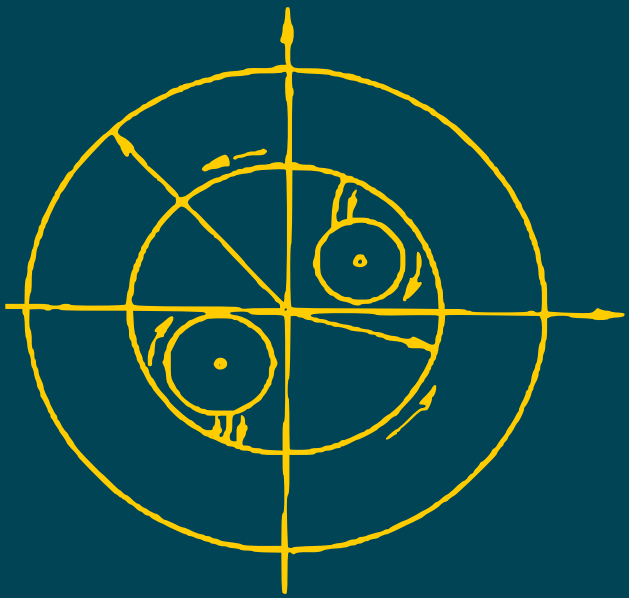
G. Chilov

Analyse Mathématique

Fonctions d'une Variable

3e Partie



Editions Mir Moscou

G. CHILOV

ANALYSE MATHÉMATIQUE

FONCTIONS D`UNE VARIABLE

TOMO II

ИЗДАТЕЛЬСТВО
МИР

Г. Е. ШИЛОВ

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

ФУНКЦИИ ОДНОГО ПЕРЕМЕННОГО

Часть 3

ИЗДАТЕЛЬСТВО «НАУКА» МОСКВА

G. CHILOV

Analyse mathématique

FONCTIONS D'UNE VARIABLE

3e partie

Deuxième édition

Editions Mir • Moscou

Réimpression 1978

Traduit du russe par
Vitali Kharine

На французском языке

© Traduction française·Editions Mir 1973

Avant-propos

La troisième partie du livre « Fonctions d'une variable » es basée sur les mêmes principes que les deux premières précédemment parues. Ils sont exprimés dans l'avant-propos du premier tome. La numérotation des chapitres du présent volume (12-16) continue celle du précédent (1-11).

Dans la troisième partie, le rôle principal appartient au chapitre 12 « Structures fondamentales de l'analyse » où l'on considère les espaces vectoriels, les espaces métriques (contrairement au chapitre 3 de la première partie, ici ce sont des espaces fonctionnels, et non pas des ensembles de points dans un espace de dimension finie, qui en servent de modèles), les espaces normés, les algèbres normées et, enfin, les espaces hilbertiens. Les algèbres normées sont appliquées à la théorie des opérateurs linéaires dans un espace normé ; en particulier, le « calcul opérationnel » des fonctions analytiques dans une algèbre normée, appliqué à l'algèbre des opérateurs linéaires, conduit à des théorèmes du genre de l'alternative de Fredholm. L'étude de l'espace vectoriel normé des suites bornées et celle des fonctionnelles sur cet espace sont liées aux notions de limite généralisée et de sommation généralisée des séries.

Dans le chapitre 13 « Equations différentielles », on établit les théorèmes principaux sur les solutions des équations différentielles ordinaires pour les fonctions à valeurs dans un espace normé. La solution d'une équation linéaire à coefficient opératoriel constant s'exprime par l'exponentielle d'un opérateur ; en l'explicitant nous obtenons les formules pour les solutions d'une équation linéaire à coefficients constants, d'un système d'équations de ce type et d'une équation d'ordre supérieur. Pour une équation linéaire à coefficient opératoriel variable, on construit la méthode de variation de la constante.

C'est essentiellement les séries de Fourier que l'on étudie dans le chapitre 14 « Développements orthogonaux » ; on considère de divers types de convergence et de sommabilité de ces séries.

Le chapitre 15 « Transformation de Fourier », parallèlement à la théorie réelle ordinaire, traite des problèmes liés au domaine complexe, en particulier à la transformation de Laplace.

Dans le chapitre 16 « Courbes gauches », nous exposons la théorie de la courbure dans un espace à plusieurs dimensions.

Comme dans les deux premières parties, l'exposé est accompagné d'exercices. On trouve les réponses et les indications correspondantes à la fin du livre.

L'auteur

Ainsi, avec ces indispensables correctifs, peut-on mieux prendre conscience de la vie interne de la mathématique, de ce qui fait à la fois son unité et sa diversité; telle une grande cité dont les faubourgs ne cessent de progresser, de façon quelque peu chaotique, sur le terrain environnant, tandis que le centre se reconstruit périodiquement, chaque fois suivant un plan plus clair et une ordonnance plus majestueuse, jetant à bas les vieux quartiers et leurs dédales de ruelles, pour lancer vers périphérie des avenues toujours plus directes, plus larges et plus commodes.

N. Bourbaki, L'architecture des mathématiques (1938)

TROISIEME PARTIE

Chapitres choisis de l'analyse moderne

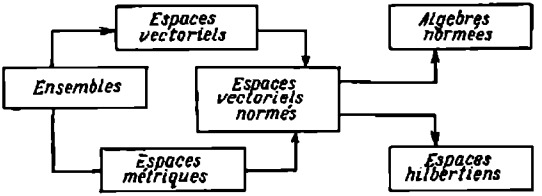
Structures fondamentales de l'analyse

H i l b a r t. C'est donc celui-là? Mais bien sûr que je m'en souviens, il était mon élève dans le temps. Après, il est devenu poète: évidemment, il n'avait pas assez de fantaisie pour s'occuper des mathématiques.

Nous avons déjà parlé de structures mathématiques (§ 2.5). Disons-en encore quelques mots visant cette fois les structures qui se présentent dans l'analyse. Les objets de l'analyse mathématique sont les nombres, les fonctions et les opérations sur ces nombres et fonctions. Du point de vue le plus général, les liens qui existent entre ces objets sont décrits par la théorie des ensembles. En effet, nombres et fonctions forment des ensembles variés; les relations d'inclusion, les opérations de réunion, intersection, passage au complémentaire permettent de décrire certaines propriétés générales de ces ensembles. Nous arrivons à des structures fondamentales de l'analyse en imposant aux ensembles les conditions supplémentaires mises sous la forme d'un système d'axiomes correspondant à certaines propriétés ou opérations que l'on utilise dans l'analyse mathématique classique. C'est ainsi qu'apparaissent les structures mathématiques suivantes: espace vectoriel où l'on axiomatise les opérations linéaires d'addition des éléments et de multiplication d'un élément par un nombre; espace métrique où, à l'aide de la notion de distance, on axiomatise l'opération de passage à la limite; espace vectoriel normé (« de Banach ») où l'on considère les opérations linéaires ainsi que le passage à la limite; algèbre normée où l'on ajoute aux opérations mentionnées celle de multiplication des éléments; espace hilbertien où l'on axiomatise la notion de produit scalaire, ce qui permet d'opérer non seulement avec les longueurs de vecteurs mais aussi avec les angles qu'ils forment; onfin, lorsqu'on demande que le nombre de dimensions soit fini, on aboutit aux espaces vectoriels affine (i.e. non métrisé), normé et hilbertien (ou euclidien) de dimension finie. A part les structures fondamentales mentionnées, il existe une quantité de structures intermédiaires dont nous ne parlons pas pour l'instant bien qu'elles

soient très importantes (espaces topologiques, espaces partiellement ordonnés, etc.).

Voici le schéma des structures fondamentales que nous allons étudier plus on détail :



Chaque flèche désigne une déduction, i.e. un passage d'une notion générale à une notion spéciale.

§ 12.1. Espaces vectoriels *)

12.11. On construit l'axiomatique de l'espace vectoriel en partant des propriétés de l'espace réel n-dimensionnel R_n (2.61) mais sans tenir compte de la notation utilisant les coordonnées et en remplaçant le corps R des nombres réels par un corps quelconque K (1.22). Notamment, un *espace vectoriel* K sur le corps K est un ensemble des objets x, y, ... appelés *vecteurs*, pour lesquels on a établi les opérations d'addition et de multiplication par les nombres (du corps K) de sorte que les axiomes suivants soient satisfaits :

a. $x + y = y + x$ quels que soient x et y de K.

b. $(x + y) + z = x + (y + z)$ quels quo soient x, y, z de K.

c. Il existe dans K un vecteur désigné par 0 (*vecteur nul*) tel que $x + 0 = x$ pour tout $x \in$ K.

d. Pour tout $x \in$ K il y a un élément $y \in$ K dit *opposé* de x tel que $x + y = 0$.

e. $\alpha(x + y) = \alpha x + \alpha y$ quels que soient x, $y \in$ K et $\alpha \in K$.

f. $(\alpha + \beta)\, x = \alpha x + \beta x$ quels que soient $x \in$ K et α et β de K.

g. $1 \cdot x = x$ pour tout $x \in$ K.

h. $\alpha(\beta x) = (\alpha\beta)\, x$ quels que soient $x \in$ K et α et β de K.

Lorsque le corps K est le corps R des nombres réels, l'espace K est appelé espace vectoriel *réel* et est désigné par R. Lorsque le corps K est le corps C des nombres complexes, l'espace K est appelé espace vectoriel *complexe* et désigné par C.

12.12. Les axiomes de l'addition *a-d* répètent les axiomes de 1.21 pour les nombres réels. C'est pourquoi, dans tout espace vectoriel,

*) Pour plus de détails, cf. [14].

les corollaires que nous avons tirés dans § 1.3 des axiomes de l'addition des nombres réels sont valables, à savoir: unicité du zéro, unicité de l'opposé pour tout $x \in K$, existence et unicité de la solution de l'équation $a + x = b$, ce qui garantit la possibilité d'une définition correcte de l'opération de soustraction.

L'opération de multiplication des éléments d'un espace vectoriel n'est pas définie, et la ressemblance des axiomes *e-h* avec certains axiomes de la multiplication des nombres réels cités dans 1.22 est trompeuse. Pour cette raison, quelques-uns seulement des théorèmes de § 1.4 sont valables pour les espaces vectoriels. Restent justes, sans que la démonstration change tant soit peu sérieusement, les propositions suivantes:

a (analogue de 1.47*a*). *Pour tout* $x \in K$ *on a l'égalité* $0 \cdot x = 0$ (ici 0 dans le second membre est le vecteur nul et dans le premier le nombre 0 du corps K).

b (analogue de 1.47*b*). *Si* $\alpha x = 0$, *alors ou bien* $\alpha = 0$, *ou bien* $x = 0$.

En effet, si $\alpha \neq 0$, on a d'après 12.11 *g-h*:

$$x = \frac{1}{\alpha} \alpha x = \frac{1}{\alpha} \cdot 0 = 0.$$

c (analogue de 1.49). *Pour tout* $x \in K$ *l'égalité* $-x = (-1)\, x$ *a lieu.*

12.13. Exemples d'espaces vectoriels. Signalons quatre types d'espaces sur le corps R des nombres réels:

a. Nombres réels eux-mêmes avec les opérations habituelles.

b. Espace réel R_n de dimension n (§ 2.6).

c. Espace $R(E)$ de toutes les fonctions (à valeurs réelles) définies sur un ensemble E, avec les opérations habituelles (pour les fonctions numériques) d'addition et de multiplication par les nombres réels (4.31*b*).

d. Espace $\mathrm{R}(E)$ de toutes les fonctions à valeurs vectorielles (d'un espace réel R) avec les opérations d'addition et de multiplication par les nombres réels définies d'une façon naturelle pour les fonctions à valeurs vectorielles:

$$(x + y)(t) = x(t) + y(t), \quad (\alpha x)(t) = \alpha x(t).$$

Chacun de ces exemples sauf le premier est une généralisation du précédent.

En remplaçant dans ces exemples le corps des nombres réels par un corps quelconque K, on obtient quatre exemples d'espaces sur le corps K:

e. Corps K lui-même.

f. Espace n-dimensionnel K_n sur le corps K, formé de tous les complexes $(\alpha_1, \ldots, \alpha_n)$ composés chacun de n éléments du corps K

avec les opérations d'addition et de multiplication par les nombres introduites d'après les règles:

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n),$$
$$\beta\,(\alpha_1, \ldots, \alpha_n) = (\beta\alpha_1, \ldots, \beta\alpha_n).$$

g. Espace $K\,(E)$ de toutes les fonctions à valeurs dans le corps K définies sur l'ensemble E avec les opérations habituelles (pour les fonctions) d'addition et de multiplication par un nombre.

h. Espace K (E) de toutes les fonctions à valeurs vectorielles (d'un espace K) définies sur l'ensemble E et munies des opérations naturelles (pour les fonctions vectorielles) d'addition et de multiplication par un nombre.

L'espace suivant ne figure pas sur la liste ci-dessus mais est d'une importance primordiale dans l'analyse:

i. Espace $R^s\,(M)$ de toutes les fonctions réelles continues définies sur un espace métrique M *).

Cet exemple ne peut être généralisé au cas des fonctions à valeurs dans un corps quelconque K puisque, pour de telles fonctions, la notion de continuité n'est en général pas définie (la continuité d'une fonction suppose une métrique dans l'espace où la fonction prend ses valeurs; or on n'a introduit aucune métrique dans un corps arbitraire K).

Pour le moment, on ne peut étendre l'exemple *i* qu'au cas des fonctions à valeurs dans l'espace réel n-dimensionnel R_n où l'on a aussi bien les opérations linéaires que la notion de continuité (5.81):

j. Espace $R_n^s\,(M)$ de toutes les fonctions continues définies sur un espace métrique M, à valeurs dans l'espace réel R_n de dimension n.

k. Un cas particulier de l'exemple *j* qui mérite d'être signalé est l'espace $C^s\,(M)$ de toutes les fonctions continues sur un espace métrique M, à valeurs complexes.

Une restriction utile de l'exemple *i* sera considérée dans 12.23*b*.

12.14.a. Les vecteurs $x_1, \ldots, x_n$ d'un espace vectoriel K sont dits *linéairement dépendants* s'il existe dans le corps K les constantes $\alpha_1, \ldots, \alpha_n$ qui ne sont pas toutes nulles et pour lesquelles

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0. \tag{1}$$

Mais si l'égalité (1) implique $\alpha_1 = \ldots = \alpha_n = 0$, alors les vecteurs $x_1, \ldots, x_n$ sont dits *linéairement indépendants.*

b. Un espace vectoriel K est dit *à n dimensions* (ou de dimension n) s'il y existe n vecteurs linéairement indépendants mais si n'importe quels $n + 1$ vecteurs sont linéairement dépendants. Si, dans un espace K, pour tout $n = 1, 2, \ldots$, il existe n vecteurs

*) L'indice supérieur veut dire « continu » (stetig, en allemand).

linéairement indépendants, alors l'espace K est dit *de dimension infinie*.

c. On appelle *base* d'un espace n-dimensionnel K tout ensemble de ses n vecteurs linéairement indépendants. Si $f_1, \ldots, f_n$ est une base et x un vecteur quelconque de l'espace K, alors $n+1$ vecteurs $x, f_1, \ldots, f_n$ sont déjà linéairement dépendants, et il existe donc les nombres $\alpha_0, \alpha_1, \ldots, \alpha_n$ de K, pas tous nuls, tels que

$$\alpha_0 x + \alpha_1 f_1 + \ldots + \alpha_n f_n = 0.$$

De plus $\alpha_0 \neq 0$, sinon les vecteurs $f_1, \ldots, f_n$ s'avéreraient linéairement dépendants. En divisant par α_0 et en posant $\beta_j = -\alpha_j/\alpha_0$ $(j = 1, \ldots, n)$, on obtient une *décomposition du vecteur x suivant la base* $f_1, \ldots, f_n$:

$$x = \beta_1 f_1 + \ldots + \beta_n f_n.$$

Une telle décomposition est *unique* (sinon les vecteurs $f_1, \ldots, f_n$ seraient linéairement dépendants).

d. L'espace réel n-dimensionnel R_n (2.61) est bien un espace vectoriel n-dimensionnel au sens de la définition donnée. Notamment, les vecteurs

$$e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$$

sont, évidemment, linéairement indépendants. Or, n'importe quels $n+1$ vecteurs

$$y_1 = (x_1^{(1)}, \ldots, x_n^{(1)}),$$
$$\ldots\ldots\ldots\ldots$$
$$y_{n+1} = (x_1^{(n+1)}, \ldots, x_n^{(n+1)})$$

sont linéairement dépendants, ce qu'on a déjà vu dans 2.64.

De la même façon, l'espace K_n (12.13*f*) est de dimension n au sens de ladite définition.

e. Soit Ω un ensemble infini sur la droite numérique $-\infty < x < \infty$. Désignons par $P(\Omega)$ l'espace vectoriel de tous les polynômes $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ (de tous les degrés) définis sur Ω, à coefficients d'un corps quelconque K et avec les opérations habituelles; l'espace $P(\Omega)$ est un espace vectoriel sur le corps K. Montrons que, *pour tout n, les fonctions* $1, x, \ldots, x^n$ *sont linéairement indépendantes*. Supposons que sur Ω on ait l'égalité

$$\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n \equiv 0.$$

En substituant successivement à x les valeurs (différentes) $x_0, x_1, x_2, \ldots, x_n$ (de Ω), on obtient le système d'équations par

rapport à $\alpha_0, \ldots, \alpha_n$:

$$\left.\begin{aligned} &\alpha_0 + \alpha_1 x_0 + \ldots + \alpha_n x_0^n = 0, \\ &\alpha_0 + \alpha_1 x_1 + \ldots + \alpha_n x_1^n = 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\alpha_0 + \alpha_1 x_n + \ldots + \alpha_n x_n^n = 0 \end{aligned}\right\}$$

dont le déterminant n'est pas nul (déterminant de Vandermonde). D'où $\alpha_0 = \alpha_1 = \ldots = \alpha_n = 0$, ce qu'il fallait démontrer.

D'après la définition donnée dans *b*, *l'espace* $P(\Omega)$ *est de dimension infinie.*

f. Montrons que *l'espace* $R^s(M)$ $(C^1(M))$ *de toutes les fonctions réelles (complexes) continues sur un espace métrique infini* M *est de dimension infinie.*

Pour tout $n = 1, 2, \ldots$, nous indiquons n fonctions linéairement indépendantes dans l'espace $R^s(M)$. Soient $t_1, \ldots, t_n$ des points différents de l'espace M et $d = \min_{j,\,k} \rho(t_j, t_k)$. Considérons une fonction continue $y = \varphi(x)$ de l'argument réel qui vaut 1 pour $x = 0$ et 0 pour $|x| \geqslant d$. La fonction $\rho(t_j, t)$ est continue par rapport à t (5.12*b*), donc $x_j(t) = \varphi[\rho(t_j, t)]$ est aussi continue par rapport à t (5.15). Par construction, la fonction $x_j(t)$ vaut 1 pour $t = t_j$ et 0 pour $t = t_k$, $k \neq j$. Supposons qu'il existe la relation

$$\alpha_1 x_1(t) + \ldots + \alpha_n x_n(t) = 0 \quad \text{sur } M.$$

En y posant $t = t_j$ on obtient $\alpha_j = 0$ $(j = 1, \ldots, n)$, d'où l'indépendance linéaire des fonctions $x_j(t)$.

g. Un sous-ensemble $E \subset K$ est appelé *sous-espace* d'un espace K si $x \in E$, $y \in E$ impliquent $x + y \in E$, $\alpha x \in E$ pour tout nombre $\alpha \in K$.

Dans tout espace vectoriel K il existe deux sous-espaces impropres : le premier est formé d'un seul élément 0 et appelé sous-espace *nul*, le second est identique à tout l'espace K. Tous les autres sous-espaces de K sont appelés sous-espaces *propres*.

h. Sommes directes. On dit qu'un espace K est *somme directe* de ses sous-espaces $L_1, \ldots, L_n$ lorsque, pour tout $x \in K$, il existe une décomposition

$$x = x_1 + \ldots + x_n, \quad x_1 \in L_1, \ldots, x_n \in L_n,$$

et que cette décomposition est unique, c'est-à-dire que

$$x = x_1 + \ldots + x_n = y_1 + \ldots + y_n, \quad x_j \in L_j, \ y_j \in L_j, \quad (2)$$
$$j = 1, \ldots, n,$$

implique $x_1 = y_1, \ldots, x_n = y_n$.

La condition (2) d'unicité de la décomposition de tout élément x peut être remplacée par la condition plus simple d'unicité

de la décomposition du zéro: s'il y a une décomposition

$$0 = x_1 + \ldots + x_m, \quad x_1 \in L_1, \ldots, x_m \in L_m, \tag{3}$$

alors $x_1 = \ldots = x_m = 0$.

Ainsi, l'espace R_n est la somme directe de n sous-espaces unidimensionnels engendrés par n vecteurs linéairement indépendants quelconques. D'ailleurs, on peut mettre l'espace R_n sous forme de somme directe de sous-espaces non unidimensionnels, et cela de diverses manières. En général, pour tout sous-espace $L \subset R_n$, il existe un autre sous-espace $M \subset R_n$ tel que la somme directe de M et L donne tout l'espace R_n.

Si un espace vectoriel K est mis sous forme de somme directe des sous-espaces $L_1, \ldots, L_m$, *alors tout couple de ces derniers n'a qu'un seul vecteur commun, à savoir le vecteur nul* [14; 2.45].

1. Espace quotient. Deux éléments $x \in K$, $y \in K$ sont dits *équivalents par rapport à un sous-espace* $L \subset K$ si $x - y \in L$. La relation d'équivalence est désignée par $x \underset{L}{\sim} y$ ou, en abrégé, $x \sim y$.

L'ensemble X de tous les éléments y équivalents à un élément donné x est appelé *classe d'équivalence selon le sous-ensemble* L ou *classe* tout court. La classe X contient l'élément x lui-même; deux éléments quelconques d'une même classe sont équivalents; enfin, si $z \notin X$, alors z n'est équivalent à aucun élément $y \in X$. Par conséquent, deux classes ou bien n'ont aucun élément commun ou bien se confondent.

Tout l'espace K représente la réunion des classes deux à deux disjointes X, Y, ... L'ensemble de ces classes est désigné par K/L. Dans l'ensemble K/L, on définit les opérations linéaires de la façon suivante. Soient X et Y des classes, α et β des nombres; nous voulons définir la classe $Z = \alpha X + \beta Y$. Pour le faire, choisissons arbitrairement les éléments $x \in X$ et $y \in Y$ et trouvons la classe Z qui contient l'élément $z = \alpha x + \beta y$, désignons cette dernière par $\alpha X + \beta Y$. On démontre qu'elle est définie d'une façon unique et que les opérations ainsi introduites satisfont aux axiomes 12.11. Le zéro de l'espace K/L est la classe qui contient 0 de l'espace K et qui est donc identique au sous-espace L. L'opposé de la classe X est la classe formée des éléments opposés de ceux de la classe X. Les démonstrations de toutes ces propositions sont données dans [14; 2.48].

L'espace K/L ainsi construit est appelé *espace quotient de l'espace* K *par le sous-espace* L.

j. Morphismes des espaces vectoriels. Soient X et Y deux espaces vectoriels sur un même corps K. Une application $y = \omega(x)$ de l'espace X dans l'espace Y est appelée *morphisme* (homomorphisme, opérateur linéaire) *de l'espace* X *dans l'espace* Y si, pour deux éléments quelconques x_1, x_2 de l'espace X

et pour deux nombres quelconques α_1, α_2 de K, l'égalité

$$\omega(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \omega(x_1) + \alpha_2 \omega(x_2)$$

est vérifiée.

Si un morphisme ω est une application de l'espace X sur tout l'espace Y, on l'appelle *épimorphisme*. Si un morphisme ω est une application non nécessairement sur tout Y mais injective, de sorte que $x_1 \neq x_2$ implique $\omega(x_1) \neq \omega(x_2)$, alors il est appelé *monomorphisme*. Un morphisme ω qui est à la fois épimorphisme et monomorphisme, i.e. une application injective de l'espace X sur tout l'espace Y et qui conserve les opérations linéaires, s'appelle *isomorphisme* (en conformité avec la définition générale de l'isomorphisme des structures 2.52). Un morphisme est souvent désigné comme suit :

$$\omega : X \to Y.$$

Si X est un sous-espace d'un espace Y, alors l'application ω qui à tout élément $x \in X$ fait correspondre lui-même en tant qu'élément de l'espace Y est un monomorphisme $\omega : X \to Y$, et l'application ω' qui à tout élément $x \in Y$ fait correspondre la classe $U \in Y/X$ comprenant cet élément x est un épimorphisme $\omega' : Y \to Y/X$.

Théorème. *Tout espace abstrait n-dimensionnel* K_n *sur un corps K est isomorphe à l'espace n-dimensionnel* K_n.

Démonstration. Soit $f_1, \ldots, f_n$ un système de n vecteurs linéairement indépendants de l'espace K_n. Pour tout $x \in K_n$, il existe une représentation $x = \alpha_1 f_1 + \ldots + \alpha_n f_n$ et une seule (*c*). Faisons correspondre au vecteur x le vecteur $y = (\alpha_1, \ldots, \alpha_n) \in K_n$. Nous obtenons une application bijective $\omega : K_n \to K_n$ qui, comme c'est facile à vérifier, conserve les opérations linéaires, i.e. est un isomorphisme.

Exemple. R_n/R_m $(n > m)$ est isomorphe à R_{n-m}.

k. Produits cartésiens. Si X et Y sont deux espaces vectoriels, on peut former leur produit cartésien P (X, Y) (2.82) de tous les couples possibles (x, y) $x \in X$, $y \in Y$. Dans le produit cartésien, on introduit les opérations linéaires « par coordonnées » :

$$\alpha_1(x_1, y_1) + \alpha_2(x_2, y_2) = (\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2).$$

On vérifie aisément que les axiomes 12.11 sont remplis. Il est évident que l'espace P (X, Y) contient deux sous-espaces

$$X^* = \{(x, y) : y = 0\}, \quad Y^* = \{(x, y) : x = 0\},$$

qui sont isomorphes (*j*) respectivement aux espaces X et Y. De plus, pour tout élément $(x, y) \in P(X, Y)$, on a :

$$(x, y) = (x, 0) + (0, y).$$

La dernière décomposition de l'élément (x, y) en deux termes appartenant respectivement à X^* et Y^* est unique (en vertu de la définition de l'addition dans $P(X, Y)$ et de l'égalité des éléments dans $P(X, Y)$). Ainsi, le produit cartésien de deux espaces X et Y est la somme directe de ses sous-espaces X^* et Y^* isomorphes respectivement à X et Y.

12.15. Opérateurs linéaires.

a. Dans les raisonnements analytiques, les morphismes des espaces vectoriels sont souvent appelés *opérateurs linéatres*. Ainsi, *un opérateur linéaire d'un espace vectoriel* X *dans un espace vectoriel* Y est une application $A : X \to Y$ vérifiant la condition

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2$$

pour tous x_1 et x_2 de X et tous α_1 et α_2 de K. Si $X = Y$, il s'agit d'un opérateur linéaire A *dans l'espace* X.

b. L'opérateur qui à tout vecteur $x \in X$ fait correspondre le vecteur nul de l'espace Y est, évidemment, un opérateur linéaire de X dans Y ; il s'appelle *opérateur nul*.

c. L'opérateur qui à tout vecteur $x \in X$ fait correspondre le même vecteur x est un opérateur linéaire dans X ; cet opérateur est appelé *opérateur unité* et désigné par E.

d. Lorsque l'espace Y est unidimensionnel, tout opérateur linéaire A est appelé *fonctionnelle linéaire*. Ce terme est surtout employé dans le cas où X est un espace de dimension infinie ; en cas de dimension finie on dira plutôt « fonction linéaire ».

e. S'il y a deux opérateurs linéaires A_1 et A_2 d'un espace X dans un espace Y, on peut définir leur *somme* $A_1 + A_2$ et le *produit de l'opérateur* A_1 *par un nombre* α suivant les règles :

$$(A_1 + A_2)\, x = A_1 x + A_2 x,$$

$$(\alpha A_1)\, x = \alpha\, (A_1 x) ;$$

dans les deux cas, on obtient toujours les opérateurs linéaires de X dans Y.

f. Il est aisé de constater que les opérations d'addition et de multiplication des opérateurs par un nombre satisfont aux mêmes axiomes 12.11 qui régissent les opérations dans un espace vectoriel. Ainsi, *l'ensemble* L (X, Y) *de tous les opérateurs linéaires d'un espace vectoriel* X *dans un espace vectoriel* Y *est lui-même un espace vectoriel*. L'espace L (X, Y) a pour zéro l'opérateur nul (*b*).

g. Multiplication des opérateurs. Si B est un opérateur linéaire d'un espace X dans un espace Y et A un opérateur linéaire de l'espace Y dans un espace Z (tous les espaces étant sur un même corps K), alors l'opérateur $P = A \cdot B$, ou plus brièvement AB, est défini comme opérateur de X dans Z d'après la formule

$$Px = (AB)\, x = A\, (Bx)$$

(i.e. d'abord l'opérateur B agit sur un vecteur $x \in X$, ensuite l'opérateur A agit sur le résultat Bx qui appartient à l'espace Y). L'opérateur obtenu $P = AB$ est un opérateur linéaire de X dans Z. Les égalités suivantes ont lieu :

$$\alpha (AB) = (\alpha A) B = A (\alpha B),$$
$$A (B_1 + B_2) = AB_1 + AB_2,$$
$$(A_1 + A_2) B = A_1B + A_2B,$$
$$A (BC) = (AB) C,$$

qui expriment les lois associatives et distributives pour la multiplication des opérateurs. Dans ces égalités, $\alpha \in K$ est un nombre arbitraire; A, A_1, A_2 sont des opérateurs de l'espace Y dans l'espace Z; B, B_1, B_2 des opérateurs de l'espace X dans l'espace Y ; C est un opérateur d'un espace W dans l'espace X ; les deux membres de trois premières égalités sont des opérateurs de X dans Z, ceux de la dernière égalité des opérateurs de W dans Z. Si E_X (E_Y) est l'opérateur unité dans l'espace X (Y), alors, pour tout opérateur B de X dans Y, a lieu encore une égalité :

$$E_Y B = B E_X = B.$$

Les opérateurs dans X peuvent être multipliés l'un par l'autre dans n'importe quel ordre ; dans les deux cas possibles, le résultat est un opérateur dans X. Mais cette multiplication n'est en général pas commutative, de sorte que, pour certains couples A,B d'opérateurs, on a $AB \neq BA$. Toujours dans le cas d'un opérateur A dans X, on définit ses puissances :

$$A^0 = E_X,\ A^1 = A,\ A^2 = A \cdot A,\ \ldots,\ A^{k+1} = A^k \cdot A,\ \ldots$$

h. A tout polynôme $p(\lambda)$ à coefficients du corps K :

$$p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k \tag{1}$$

et à tout opérateur A dans l'espace X, on peut faire correspondre un « polynôme d'un opérateur »

$$p(A) = \sum_{k=0}^{n} a_k A^k, \tag{2}$$

qui est encore un opérateur linéaire dans l'espace X ; la somme et le produit de deux polynômes de la forme (2) correspondent à la somme et au produit des polynômes correspondants de la forme (1).

i. Soient A un opérateur d'un espace Y dans un espace X et B un opérateur de X dans Y. Alors, si $AB = E_X$, l'opérateur A est appelé l'*inverse à gauche* de l'opérateur B et l'opérateur B l'*inverse à droite* de l'opérateur A. Si l opérateur B agit dans X, alors, parmi

tous les opérateurs dans X, il peut y avoir son inverse à gauche ainsi que son inverse à droite. Lorsqu'un tel opérateur possède un inverse à gauche A et un inverse à droite C, ces opérateurs inverses se confondent :

$$A = AE_X = A(BC) = (AB)C = E_XC = C.$$

Il résulte de l'égalité écrite que, dans ce cas, *n'importe quel inverse à gauche (à droite) de l'opérateur* B *est identique à* $A = C$. Cet opérateur $A = C$ défini d'une façon unique s'appelle opérateur *inverse* de l'opérateur B et est désigné par B^{-1}.

Dans les espaces de dimension infinie, il y a des opérateurs admettant un inverse à gauche (même un ensemble infini d'inverses à gauche différents) et ne possédant aucun inverse à droite ou vice versa.

j. Soit un opérateur A dans un espace X. Un sous-espace $X' \subset X$ est dit *invariant par l'opérateur* A si $x \in X'$ implique $Ax \in X'$.

k. Un vecteur $f \in X$ non nul est appelé *vecteur propre d'un opérateur* A dans l'espace X si

$$Af = \lambda f \quad (\lambda \in K);$$

le nombre λ est appelé *valeur propre de l'opérateur* A *associée au vecteur propre* f. Il est évident qu'un vecteur propre f engendre le sous-espace invariant unidimensionnel formé de tous les vecteurs αf, $\alpha \in K$.

Toute combinaison linéaire de vecteurs propres de l'opérateur A associés à une même valeur propre λ est, évidemment, encore un vecteur propre de l'opérateur A pour la même valeur propre λ. Il en découle que *l'ensemble de tous les vecteurs propres de l'opérateur* A *pour une valeur propre donnée* λ *est un sous-espace dans l'espace* X ; il est appelé *sous-espace propre de l'opérateur* A *associé à la valeur propre* λ.

l. *Les vecteurs propres* $f_1, \ldots, f_n$ *de l'opérateur* A *associés respectivement aux valeurs propres différentes* $\lambda_1, \ldots, \lambda_n$ *sont linéairement indépendants*. En effet, une fois supposée la dépendance linéaire de n vecteurs propres, on a $\alpha_1 f_1 + \ldots + \alpha_n f_n = 0$ et, en appliquant l'opérateur A et en éliminant l'un des vecteurs, on peut passer à la dépendance linéaire d'un nombre inférieur de vecteurs propres, ce qui permet d'appliquer le raisonnement par récurrence.

12.16. Exemples d'opérateurs linéaires dans les espaces concrets.

a. Soit $A = \| a_{jk} \|$ une $m \times n$-matrice (i.e. une matrice à m lignes et n colonnes) formée d'éléments d'un corps K. Choisissons une base $e_1, \ldots, e_n$ dans un espace n-dimensionnel K_n et une base $f_1, \ldots, f_m$ dans un espace m-dimensionnel K_m. A tout vecteur $x = \sum_{k=1}^{n} \xi_k e_k \in K_n$ faisons correspondre un vecteur $y = \sum_{j=1}^{m} \eta_j f_j \in K_m$

selon la règle

$$\eta_j = \sum_{k=1}^{n} a_{jk}\xi_k, \quad j = 1, \ldots, m.$$

On obtient ainsi un opérateur linéaire de l'espace K_n dans l'espace K_m.

b. Dans le cas du continu, l'opérateur

$$y(s) = Ax(s) = \int_a^b A(s, t)\, x(t)\, dt \tag{1}$$

est analogue à celui de l'exemple *a*. Ici $x(t)$ est un élément de l'espace $R^s(a, b)$, $A(s, t)$ une fonction réelle de deux variables définie pour $a \leqslant t \leqslant b$, $c \leqslant s \leqslant d$; $y(s)$ est une fonction définie sur le segment $c \leqslant s \leqslant d$. On vérifie aisément que la fonction $y(s)$ est continue sur $[c, d]$ dès que la fonction $A(s, t)$ est continue sur le rectangle $a \leqslant t \leqslant b$, $c \leqslant s \leqslant d$. Dans ce cas l'opérateur A est un opérateur linéaire de l'espace $R^s(a, b)$ dans l'espace $R^s(c, d)$.

L'opérateur (1) est appelé *opérateur intégral de Fredholm*. Les opérateurs de Fredholm sont considérés plus en détail dans 12.98.

c. Un cas spécial de l'opérateur *b* est représenté par l'opérateur d'intégration

$$Ix(t) = \int_a^t x(\tau)\, d\tau, \quad a \leqslant t \leqslant b,$$

dans l'espace $R^s(a, b)$.

d. L'expression

$$F(x) = \int_a^b f(\tau)\, x(\tau)\, d\tau$$

avec une fonction $f(t)$ (continue) fixe fournit un exemple de fonctionnelle linéaire définie dans l'espace $R^s(a, b)$.

12.17. Opérateurs linéaires dans les espaces de dimension finie.

a. Donnons la forme générale d'un opérateur linéaire A d'un espace *n*-dimensionnel K_n dans un espace *m*-dimensionnel K_m. Soient $e_1, \ldots, e_n$ une base dans l'espace K_n et $f_1, \ldots, f_m$ une base dans l'espace K_m. En appliquant l'opérateur A aux vecteurs $e_1, \ldots, e_n$ on trouve

$$\left.\begin{aligned} Ae_1 &= a_{11}f_1 + \ldots + a_{m1}f_m, \\ Ae_2 &= a_{12}f_1 + \ldots + a_{m2}f_m, \\ &\ldots\ldots\ldots\ldots\ldots \\ Ae_n &= a_{1n}f_1 + \ldots + a_{mn}f_m, \end{aligned}\right\} \tag{1}$$

a_{ij} étant des nombres du corps K.

Ainsi, les bases $\{e\}$ et $\{f\}$ dans les espaces K_n et K_m étant fixes, il correspond à l'opérateur A la $m \times n$-matrice

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}.$$

Ici la j-ième colonne est composée des coordonnées du vecteur Ae_j dans la base $f_1, \dots, f_m$.

Soit maintenant $x = \sum_1^n \xi_k e_k$ un vecteur quelconque de K_n et soit

$$Ax = \sum_1^m \eta_j f_j \in K_m.$$

Nous avons

$$\sum_1^m \eta_j f_j \equiv Ax = \sum_1^n \xi_k Ae_k = \sum_{k=1}^n \xi_k \sum_{j=1}^m a_{jk} f_j = \sum_{j=1}^m \Big(\sum_{k=1}^n a_{jk} \xi_k \Big) f_j,$$

d'où

$$\eta_j = \sum_{k=1}^n a_{jk} \xi_k \quad (j = 1, \dots, m). \tag{2}$$

Ainsi, c'est bien la forme générale d'un opérateur de l'espace K_n dans K_m qui a été donnée dans l'exemple 12.16*a*.

Si l'opérateur linéaire A agit dans K_n, alors $m = n$, et la matrice A est carrée.

Si l opérateur linéaire A applique K_n dans K_1 (espace unidimensionnel), alors $m = 1$, et la matrice A est de la forme

$$A = || a_1 \; a_2 \dots a_n ||.$$

Dans ce cas, l'opérateur A agit d'après la formule

$$Ax = \sum_{k=1}^n a_k \xi_k$$

(le seul vecteur de base do l'espace K_1 n'étant pas explicité) et représente une fonction linéaire.

b. Aux opérations sur les opérateurs linéaires définies dans 12.15*e-f*. il correspond les opérations analogues sur leurs matrices. Soient toujours $e_1, \dots, e_n$ une base dans un espace X et $f_1, \dots, f_m$ une base dans un espace Y. Aux opérateurs A_1 et A_2 appliquant X dans Y, il correspond dans ces bases les matrices $A_1 = || a_j^{(1)} ||$

et $A_2 = \| a_{ij}^{(2)} \|$ respectivement, de sorte que

$$A_1 e_j = \sum_{i=1}^{m} a_{ij}^{(1)} f_i, \quad A_2 e_j = \sum_{i=1}^{m} a_{ij}^{(2)} f_i \quad (j = 1, \ldots, n).$$

Alors, pour tous α_1, α_2 de K, on a :

$$(\alpha_1 A_1 + \alpha_2 A_2)\, e_j = \sum_{i=1}^{m} (\alpha_1 a_{ij}^{(1)} + \alpha_2 a_{ij}^{(2)})\, f_i,$$

c'est-à-dire que la matrice $\| \alpha_1 a_{ij}^{(1)} + \alpha_2 a_{ij}^{(2)} \|$ correspond à l'opérateur linéaire $\alpha_1 A_1 + \alpha_2 A_2$. Ainsi, les matrices associées à la somme des opérateurs et au produit d'un opérateur par un nombre s'obtiennent en additionnant « élément à élément » les matrices des opérateurs et en multipliant la matrice de l'opérateur par ce nombre.

c. Il en découle en particulier que *l'espace vectoriel* $L(K_n, K_m)$ *de tous les opérateurs linéaires d'un espace n-dimensionnel* K_n *dans un espace m-dimensionnel* K_m *est isomorphe à l'espace nm-dimensionnel* K_{nm}.

d. Construisons la matrice associée au produit de deux opérateurs. Choisissons une base $e_1, \ldots, e_n$ dans l'espace X, une base $f_1, \ldots, f_m$ dans l'espace Y et une base $g_1, \ldots, g_q$ dans l'espace Z. Supposons qu'un opérateur B de X dans Y ait la $m \times n$-matrice $B = \| b_{jk} \|$, de sorte que

$$B e_k = \sum_{j=1}^{m} b_{jk} f_j \quad (k = 1, \ldots, n),$$

et qu'un opérateur A de Y dans Z ait la $q \times m$-matrice $A = \| a_{ij} \|$, de sorte que

$$A f_j = \sum_{i=1}^{q} a_{ij} g_i \quad (j = 1, \ldots, m).$$

Pour le produit $P = AB$, on obtient

$$ABe_k = A(Be_k) = A\Big(\sum_{j=1}^{m} b_{jk} f_j\Big) = \sum_{j=1}^{m} b_{jk} A f_j =$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{q} b_{jk} a_{ij} g_i = \sum_{i=1}^{q} \Big(\sum_{j=1}^{m} a_{ij} b_{jk}\Big) g_i.$$

Par conséquent, les éléments p_{ik} de la matrice P de l'opérateur $P = AB$ sont de la forme

$$p_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk} \quad (i = 1, \ldots, q,\ k = 1, \ldots, n). \tag{3}$$

La matrice $P = \| p_{ik} \|$ obtenue à partir des matrices $A = \| a_{ij} \|$ et $B = \| b_{jk} \|$ d'après la formule (3) est appelée *produit* de la première matrice par la seconde.

Nous pouvons donc multiplier une $q \times m$-matrice par une $m \times n$-matrice, et nous avons pour résultat une $q \times n$-matrice.

Si $X = Y = Z$, alors A et B sont des $n \times n$-matrices carrées et leur produit AB est aussi une $n \times n$-matrice.

e. Soit un opérateur A qui agit dans un espace n-dimensionnel K_n. Si l'on connaît la matrice $\| a_{jk} \|$ de l'opérateur A par rapport à une base $\{e\} = (e_1, \ldots, e_n)$, on peut trouver les valeurs propres de l'opérateur A (12.15*k*) sous forme de ses racines caractéristiques, i.e. les racines de l'équation

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix} = 0. \tag{4}$$

Si λ_0 est une racine de l'équation (4), alors on peut trouver les coordonnées d'un vecteur propre associé $f = \sum_{k=1}^{n} \xi_k e_k$ qui sont solutions du système suivant d'équations linéaires homogènes:

$$\left.\begin{aligned} (a_{11}-\lambda_0)\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n &= 0, \\ a_{21}\xi_1 + (a_{22}-\lambda_0)\xi_2 + \ldots + a_{2n}\xi_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + (a_{nn}-\lambda_0)\xi_n &= 0, \end{aligned}\right\} \tag{5}$$

ce système possédant des solutions non nulles.

f. Décrivons la structura d'un opérateur linéaire quelconque dans un espace complexe ou réel K_n *).

Pour tout opérateur linéaire A dans un espace complexe C_n, cet espace admet une décomposition en somme directe des sous-espaces invariants dans chacun desquels, pour une base convenablement choisie, la matrice de l'opérateur A a la forme

$$\begin{Vmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda \end{Vmatrix} \tag{6}$$

(« case jordanienne »). La base de l'espace C_n obtenue en réunissant les bases des sous-espaces invariants mentionnés est appelée *base*

*) Voir [14; ch. 6].

jordanienne de l'opérateur A et la matrice de l'opérateur A par rapport à cette base (une matrice quasi diagonale à cases diagonales de la forme (6)) *matrice jordanienne de l'opérateur* A. Les nombres λ et les dimensions des cases jordaniennes (6) sont des invariants de l'opérateur A (i.e. ne dépendent pas du choix de la base jordanienne); les nombres λ sont des racines de l'équation (4), et les dimensions des cases jordaniennes peuvent être trouvées d'après les diviseurs élémentaires de l'opérateur A.

Pour tout opérateur linéaire A dans un espace réel R_n, cet espace admet une décomposition en somme directe des sous-espaces invariants dans chacun desquels, pour une base convenablement choisie, la matrice de l'opérateur A a ou bien la forme (6) ou bien la forme

$$\begin{Vmatrix} \sigma & \tau & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -\tau & \sigma & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \sigma & \tau & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & -\tau & \sigma & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & \sigma & \tau \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & -\tau & \sigma \end{Vmatrix} \tag{7}$$

(« case jordanienne réelle »). La base de l'espace R_n obtenue en réunissant les bases des sous-espaces invariants mentionnés est appelée *base jordanienne réelle de l opérateur* A et la matrice de l'opérateur A par rapport à cette base (une matrice quasi diagonale à cases diagonales de la forme (6) et (7)) *matrice jordanienne réelle de l'opérateur* A. Les nombres λ, σ, τ ainsi que les dimensions des cases jordaniennes (6) et (7) ne dépendent pas du choix de la base jordanienne réelle; les nombres λ et $\sigma + i\tau$ sont des racines de l'équation (4). et les dimensions des cases jordaniennes (6) et (7) se déterminent d'après les diviseurs élémentaires réels de l'opérateur A.

En particulier, si toutes les racines de l'équation (4) sont simples, la matrice jordanienne de l'opérateur A dans un espace complexe C_n prend la forme (les éléments non explicités étant nuls):

$$\begin{Vmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{Vmatrix} \tag{8}$$

Dans un espace réel, l'équation (4) possède, avec toute sa racine $\lambda = \sigma + i\tau$ non réelle, aussi la racine conjuguée $\bar{\lambda} = \sigma - i\tau$. Si toutes les racines sont simples et si l'on désigne les racines non réelles de (4) par $\sigma_1 \pm i\tau_1, \dots, \sigma_k \pm i\tau_k$ et les racines réelles par

$\lambda_{2k+1}, \ldots, \lambda_n$, alors la matrice jordanienne réelle de l'opérateur A prend la forme

$$\begin{Vmatrix} \sigma_1 & \tau_1 & & & & & \\ -\tau_1 & \sigma_1 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_k & \tau_k & & \\ & & & -\tau_k & \sigma_k & & \\ & & & & & \lambda_{2k+1} & \\ & & & & & & \ddots \\ & & & & & & & \lambda_n \end{Vmatrix} \tag{9}$$

Dans un espace complexe, la matrice jordanienne de tout opérateur possédant une matrice hermitienne ($\overline{a_{jk}} = a_{kj}$, $j, k = 1, \ldots, n$) par rapport à une base a elle aussi la forme diagonale ; dans ce cas les nombres correspondants λ_j sont réels. Dans un espace réel, c'est le cas de la matrice jordanienne réelle de tout opérateur possédant une matrice symétrique ($a_{jk} = a_{kj}$, $j, k = 1, \ldots, n$) par rapport à une base. Si, dans une base d'un espace réel, la matrice d'un opérateur A est antisymétrique ($a_{jk} = -a_{kj}$, $j, k = 1, \ldots, n$), alors la matrice jordanienne réelle de l'opérateur A prend la forme (9), tous les nombres $\sigma_1, \ldots, \sigma_k, \lambda_{2k+1}, \ldots, \lambda_n$ étant nuls.

12.18. Algèbres.

a. Un espace vectoriel U sur un corps K est appelé *algèbre* (plus précisément, algèbre sur K) si, pour les éléments $x, y, \ldots$ de U, une opération de multiplication désignée par $x \cdot y$ (ou xy) est définie de sorte que les conditions suivantes soient satisfaites :

(1) $\alpha (xy) = (\alpha x) y = x (\alpha y)$ pour tous x et y de U et tout $\alpha \in K$;

(2) $(xy) z = x (yz)$ quels que soient x, y, z de U ;

(3) $(x + y) z = xz + yz$ quels que soient x, y, z de U ;

(4) $x (y + z) = xy + xz$ quels que soient x, y, z de U.

Les conditions (1) et (2) sont appelées lois *associatives*, les conditions (3) et (4) lois *distributives*.

b. En général, la multiplication peut ne pas être commutative, de sorte que l'égalité $xy = yx$ peut être fausse pour certains couples x, y de U. Si l'égalité $xy = yx$ est juste pour n'importe quels x et y de U, l'algèbre U est dite *commutative*.

c. Un élément $e \in U$ s'appelle *unité* de l'algèbre U si, pour tout $x \in U$, les égalités $ex = xe = x$ ont lieu. Un élément $y \in U$ est appelé *inverse* d'un élément x si les égalités $xy = yx = e$ sont vérifiées.

d. Un sous-espace $V \subset U$ est appelé *sous-algèbre* de l'algèbre U si $x \in V$, $y \in V$ impliquent $xy \in V$.

e. Une sous-algèbre $J \subset U$ s'appelle *idéal à gauche* de l'algèbre U s'il découle de $x \in U$, $y \in J$ que $xy \in J$, et *idéal à droite* si $y \in J$, $z \in U$ entraînent $yz \in J$. Un idéal qui est à la fois idéal à gauche et idéal à droite est appelé *idéal bilatère* ou *idéal* tout court. Dans une algèbre commutative, il n'y a évidemment pas de distinction à faire entre un idéal, un idéal à gauche et un idéal à droite.

Dans toute algèbre U, il existe deux idéaux triviaux : l'un composé d'un seul élément nul et appelé *idéal nul*, et l'autre qui est identique à toute l'algèbre U. Tous les autres idéaux sont appelés *idéaux propres*.

f. Dans l'espace quotient U/J d'une algèbre U par son idéal J, on peut définir, pour les classes X, Y, . . ., non seulement les opérations linéaires (comme dans 12.14*i*), mais encore une multiplication. A savoir, étant donnés deux classes X, Y et les éléments $x \in X$, $y \in Y$ choisis arbitrairement, on définit le produit XY comme classe contenant le produit xy. On démontre que cette définition est correcte (i.e. la classe XY ne dépend pas du choix des éléments $x \in X$, $y \in Y$) et que l'espace U/J avec l'opération de multiplication introduite devient encore une algèbre. Cette algèbre s'appelle *algèbre quotient de l'algèbre* U *par son idéal* J. Elle est commutative dès que l'algèbre U l'est.

g. Morphismes d'algèbres. Soient U et V deux algèbres sur un même corps K. Une application $\omega : U \to V$ s'appelle *morphisme* (*homomorphisme*) de l'algèbre U dans l'algèbre V si elle est un morphisme de l'espace vectoriel U dans l'espace vectoriel V (12.14*j*) et si deux éléments quelconques x_1, x_2 de l'algèbre U satisfont à la relation $\omega(x_1 x_2) = \omega(x_1) \cdot \omega(x_2)$.

Un morphisme ω s'appelle *isomorphisme* (*épimorphisme*, *monomorphisme*) de l'algèbre U dans l'algèbre V s'il est un isomorphisme (épimorphisme, monomorphisme) de l'espace U dans l'espace V.

Ainsi, l'application $\omega : U \to U/J$ qui à tout élément $x \in U$ fait correspondre la classe $X \in U/J$ qui le contient est un *épimorphisme* de l'algèbre U sur l'algèbre U/J.

h. Quel que soit un morphisme $\omega : U \to V$, l'ensemble des éléments $x \in U$ pour lesquels $\omega(x) = 0$ forme un idéal J dans l'algèbre U. Etant donné ω, définissons le morphisme $\tilde{\omega}$ de l'algèbre U/J dans l'algèbre V en faisant correspondre à une classe $X \in U/J$ l'élément $\omega(x) \in V$, où x est un élément quelconque de la classe X. Ce morphisme $\tilde{\omega}$ est un monomorphisme. Si le morphisme ω est un épimorphisme de l'algèbre U dans l'algèbre V, alors $\tilde{\omega}$ est un isomorphisme [14 ; § 6.2].

12.19. Exemples d'algèbres et de leurs morphismes.

a. L'ensemble P de tous les polynômes (de tous les degrés) en λ

$$p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$$

à coefficients d'un corps K, avec les opérations d'addition et de multiplication habituelles pour les polynômes, est évidemment une algèbre.

Cette algèbre est commutative et possède une unité.

b. L'ensemble U (G) de toutes les fonctions analytiques $f(\lambda)$ définies dans un domaine G du plan complexe forme une algèbre complexe avec les opérations d'addition et de multiplication habituelles pour les fonctions (4.72). Cette algèbre est aussi commutative et possède une unité.

Dans l'algèbre U (G), il existe un opérateur qui fait correspondre à toute fonction $f(\lambda) \in U(G)$ sa dérivée $f'(\lambda)$. Il est évidemment linéaire; notons que, pour cet opérateur, la formule de Leibniz est valable:

$$(f(\lambda)\, g(\lambda))^{(m)} = \sum_{j=0}^{m} \frac{m!}{j!\,(m-j)!} f^{(j)}(\lambda) g^{(m-j)}(\lambda). \qquad (1)$$

c. Appelons *spectre* un ensemble fini de nombres $\lambda_1, \ldots, \lambda_m$ (d'un corps K) tel qu'à tout λ_k on associe un nombre naturel r_k ($k = 1, \ldots, m$) appelé sa *multiplicité*. Appelons *corpus* et désignons par f n'importe quel ensemble de $r = r_1 + r_2 + \ldots + r_m$ nombres du corps K désignés par $f_{(j)}(\lambda_k)$ ($j = 0, \ldots, r_k - 1$, $k = 1, \ldots, m$). Enfin, désignons par F (S) l'ensemble de tous les corpus sur un spectre donné S.

Introduisons dans F (S) les opérations d'addition et de multiplication selon les règles:

$$(f + g)_{(j)}(\lambda_k) = f_{(j)}(\lambda_k) + g_{(j)}(\lambda_k),$$

$$(\alpha f)_{(j)}(\lambda_k) = \alpha f_{(j)}(\lambda_k),$$

$$(fg)_{(j)}(\lambda_k) = \sum_{i=1}^{j} \frac{j!}{i!\,(j-i)!} f_{(i)}(\lambda_k)\, g_{(j-i)}(\lambda_k)$$

$$(j = 0, \ldots, r_k - 1,\ k = 1, \ldots, m).$$

Pour $j = 0$, la dernière formule doit être remplacée par la suivante:

$$(fg)_{(0)}(\lambda_k) = f_{(0)}(\lambda_k) \cdot g_{(0)}(\lambda_k).$$

L'ensemble F (S) devient alors une algèbre de dimension r sur K.

d. Soit A un opérateur linéaire dans un espace complexe n-dimensionnel C_n. L'ensemble de tous les polynômes p (A) de l'opérn-

teur A (12.15*h*) avec les opérations naturelles d'addition et de multiplication des opérateurs représente une algèbre complexe que nous désignons par P (A). Il se trouve que cette algèbre est isomorphe à l'algèbre des corpus F (S_A) (exemple *c*), S_A étant le spectre de l opérateur A, i.e. l'ensemble de toutes les valeurs propres (différentes) $\lambda_1, \ldots, \lambda_m$ de l'opérateur A telles qu'à toute valeur λ_k on associe pour multiplicité un nombre naturel r_k égal à la plus grande dimension des cases jordaniennes de l'opérateur A (12.17*f*) qui ont le nombre λ_k sur la diagonale. Cet isomorphisme se réalise de la façon suivante: à tout corpus

$$f = \{f_{(j)}(\lambda_k),\ j = 0, \ldots, r_k - 1,\ k = 1, \ldots, m\},$$

défini sur S_A, on fait correspondre l'opérateur $f(A)$ dont la matrice par rapport à la base jordanienne de l'opérateur A a la même structure quasi diagonale que la matrice de l'opérateur A lui-même, toute $p \times p$-case quasi daigonale de l'opérateur A

$$\begin{Vmatrix} \lambda_k & 1 & 0 & \ldots & 0 \\ 0 & \lambda_k & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_k \end{Vmatrix} \quad (p \leqslant r_k) \tag{2}$$

étant remplacée par la case de mêmes dimensions:

$$\begin{Vmatrix} f_{(0)}(\lambda_k) & f_{(1)}(\lambda_k) & \frac{1}{2!} f_{(2)}(\lambda_k) & \ldots & \frac{1}{(p-1)!} f_{(p-1)}(\lambda_k) \\ 0 & f_{(0)}(\lambda_k) & f_{(1)}(\lambda_k) & \ldots & \frac{1}{(p-2)!} f_{(p-2)}(\lambda_k) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & f_{(0)}(\lambda_k) \end{Vmatrix}. \tag{3}$$

Vu de plus près, cet opérateur $f(A)$ a la forme $p(A)$, où le polynôme $p(\lambda)$ vérifie les conditions

$$p^{(j)}(\lambda_k) = f_{(j)}(\lambda_k) \quad (j = 0, \ldots, r_k - 1,\ k = 1, \ldots, m),$$

de sorte que $p^{(j)}(\lambda)$ est la dérivée *j*-ième du polynôme $p(\lambda)$.

Voir la démonstration dans [14; 6.84].

e. Soit toujours un opérateur linéaire A dans un espace complexe *n*-dimensionnel C_n, et soient $\lambda_1, \ldots, \lambda_m$ ses valeurs propres que l'on suppose appartenir toutes à un domaine *G* du plan complexe. Considérons l'application ω de l'algèbre U (*G*) des fonctions analytiques dans celle des corpus F (S_A) qui à toute fonction $f(\lambda) \in U(G)$ fait correspondre le corpus des nombres $f_{(j)}(\lambda_k) = f^{(j)}(\lambda_k)$ ($j = 0, \ldots, r_k - 1,\ k = 1, \ldots, m$), où $f^{(j)}(\lambda)$ désigne la dérivée *j*-ième de $f(\lambda)$. En vertu de la formule de Leibniz (1), l application ω ost un *morphisme* de l'algèbre U (*G*) dans l'algèbre F (S_A) ; ce morphisme est même un épimorphisme car, pour tout corpus $\{f_{(j)}(\lambda_k)\}$,

on peut trouver une fonction $f(\lambda)$ de l'algèbre U (G) (même un polynôme) pour laquelle $f^{(j)}(\lambda_k) = f_{(j)}(\lambda_k)$ $(j = 0, \ldots, r_k - 1,$ $k = 1, \ldots, m)$.

L'algèbre F (S_A) étant isomorphe, à son tour, à l'algèbre P (A) des opérateurs linéaires (exemple *d*), il existe bien un épimorphisme de l'algèbre U (G) sur l'algèbre P (A) ; compte tenu de l'exemple *d*, cet épimorphisme se réalise de la façon suivante : à toute fonction $f(\lambda) \in$ U (G) on fait correspondre l'opérateur linéaire f (A) dont la matrice par rapport à la base jordanienne de l'opérateur A a la même structure quasi diagonale que la matrice de l'opérateur A lui-même, toute case quasi diagonale (2) étant remplacée par la case de mêmes dimensions :

$$\begin{Vmatrix} f(\lambda_k) & f'(\lambda_k) & \frac{1}{2!}f''(\lambda_k) & \ldots & \frac{1}{(p-1)!}f^{(p-1)}(\lambda_k) \\ 0 & f(\lambda_k) & f'(\lambda_k) & \ldots & \frac{1}{(p-2)!}f^{(p-2)}(\lambda_k) \\ 0 & 0 & f(\lambda_k) & \ldots & \frac{1}{(p-3)!}f^{(p-3)}(\lambda_k) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & f(\lambda_k) \end{Vmatrix}$$

Ainsi, les opérateurs $e^{\lambda A}$, $\sin \lambda A$, etc. ont toujours un sens. L'application $\widetilde{\omega}$: $f(\lambda) \to f(A)$ étant un morphisme, l'égalité $f(\lambda) \cdot g(\lambda) = h(\lambda)$, où $f(\lambda)$, $g(\lambda)$, $h(\lambda)$ appartiennent à U (G), a pour conséquence $f(A) \cdot g(A) = h(A)$. Par exemple, on a toujours l'égalité :

$$e^{(\alpha+\beta)A} = e^{\alpha A} \cdot e^{\beta A}.$$

f. Un spectre S avec les $\lambda_1, \ldots, \lambda_m$ complexes (exemple *c*) est dit *symétrique* lorsque, parallèlement à tout $\lambda_k = \sigma_k + i\tau_k$ non réel, S contient le nombre conjugué $\overline{\lambda}_k = \sigma_k - i\tau_k$ de même multiplicité r_k. Un corpus $f = \{f_{(j)}(\lambda_k)\}$ sur un spectre symétrique S est dit *symétrique* si tous les nombres $f_{(j)}(\lambda_k)$ sont les complexes conjugués des nombres correspondants $f_{(j)}(\overline{\lambda}_k)$. L'ensemble de tous les corpus symétriques sur un spectre symétrique S représente (avec les opérations indiquées dans *c*) une algèbre réelle que nous désignons par $F_R(S)$.

g. Soit A un opérateur linéaire dans un espace réel n-dimensionnel R_n. L'ensemble de tous les polynômes réels de l'opérateur A forme une algèbre réelle que nous désignons par P_R (A). Il se trouve que cette algèbre est isomorphe à l'algèbre des corpus symétriques (f) sur le spectre de l'opérateur A (considéré dans le prolongement complexe *) de l'espace réel R_n) ; ce spectre est toujours symétrique.

*) Cf. [14 ; 6.61].

Quant à l'isomorphisme, il se réalise de la façon suivante. A tout corpus symétrique

$$f = \{f_{(j)}(\lambda_k),\ j = 0, \ldots, r_k - 1,\ k = 1, \ldots, m\}$$

défini sur S_A, on fait correspondre l'opérateur $f(A)$ dont la matrice par rapport à la base jordanienne réelle de l'opérateur A a la même structure quasi diagonale que la matrice jordanienne réelle de l'opérateur A, toute case quasi diagonale de l'opérateur A qui a la forme (2) (avec λ_k réel) étant remplacée par une case de la forme (3), et toute case quasi diagonale de la forme

$$\begin{Vmatrix} A_k & E & 0 & \ldots & 0 \\ 0 & A_k & E & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & A_k \end{Vmatrix}, \tag{4}$$

où les blocs A_k, E, 0 ont la forme

$$A_k = \begin{Vmatrix} \sigma_k & \tau_k \\ -\tau_k & \sigma_k \end{Vmatrix}, \quad E = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}, \quad 0 = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix},$$

par la case de mêmes dimensions:

$$\begin{Vmatrix} f_{(0)}(A_k) & f_{(1)}(A_k) & \frac{1}{2!} f_{(2)}(A_k) & \ldots & \frac{1}{(p-1)!} f_{(p-1)}(A_k) \\ 0 & f_{(0)}(A_k) & f_{(1)}(A_k) & \ldots & \frac{1}{(p-2)!} f_{(p-2)}(A_k) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & f_{(0)}(A_k) \end{Vmatrix}, \tag{5}$$

où

$$f_{(j)}(A_k) = \begin{Vmatrix} \operatorname{Re} f_{(j)}(\lambda_k) & \operatorname{Im} f_{(j)}(\lambda_k) \\ -\operatorname{Im} f_{(j)}(\lambda_k) & \operatorname{Re} f_{(j)}(\lambda_k) \end{Vmatrix}.$$

On peut prouver que l'opérateur $f(A)$ a la forme $p(A)$, le polynôme $p(\lambda)$ ayant des coefficients réels et vérifiant les conditions

$$p^{(j)}(\lambda_k) = f_{(j)}(\lambda_k) \quad (j = 1, \ldots, r_k - 1,\ k = 1, \ldots, m).$$

Voir la démonstration dans [14; 6.88].

h. Soit toujours un opérateur linéaire A dans un espace réel n-dimensionnel R_n et supposons que toutes les valeurs propres de l'opérateur A, considérées dans le prolongement complexe C_n de l'espace R_n, appartiennent à un domaine G symétrique par rapport à l'axe réel. Le morphisme ω décrit dans l'exemple e fait correspondre à toute fonction analytique réelle $f(\lambda) \in U(G)$ un corpus symétrique $f_{(j)}(\lambda_k) = f^{(j)}(\lambda_k)$ $(j = 0, \ldots, r_k - 1,\ k = 1, \ldots, m)$.

L'algèbre $F_R(S_A)$ de tous les corpus symétriques étant isomorphe à l'algèbre $P_R(A)$ de tous les polynômes réels de l'opérateur A, il existe un épimorphisme de l'algèbre $U_R(G)$ de toutes les fonctions analytiques réelles dans l'algèbre $P_R(S_A)$; cet épimorphisme, compte tenu de l'exemple *g*, se réalise de la façon suivante: à toute fonction $f(\lambda) \in U_R(G)$, on fait correspondre l'opérateur linéaire $f(A)$ dont la matrice par rapport à la base jordanienne réelle de l'opérateur A a la même structure quasi diagonale que la matrice jordanienne réelle de l'opérateur A, toute case quasi diagonale de la forme (2) (avec λ_k réel) étant remplacée par une case de la forme (3) et toute case de la forme (4) par une case de la forme (5), où

$$f_{(j)}(A_k) = \begin{Vmatrix} \operatorname{Re} f^{(j)}(\lambda_k) & \operatorname{Im} f^{(j)}(\lambda_k) \\ -\operatorname{Im} f^{(j)}(\lambda_k) & \operatorname{Re} f^{(j)}(\lambda_k) \end{Vmatrix}$$

i. L'espace vectoriel de tous les opérateurs linéaires agissant dans un espace vectoriel K forme une algèbre (avec les opérations d'addition et de multiplication habituelles pour les opérateurs) qui, en général, n'est pas commutative.

§ 12.2. Espaces métriques

12.21. A partir du chapitre 3 les espaces métriques jouent un rôle important dans notre cours. Rappelons les axiomes de l'espace métrique. Un ensemble M s'appelle *espace métrique* si, pour tout couple x, y de ses points, est défini un nombre $\rho_M(x, y)$, ou plus brièvement $\rho(x, y)$, appelé *distance* de x et y et satisfaisant aux conditions suivantes:

a. $\rho(x, y) > 0$ si $x \neq y$, $\rho(x, x) = 0$ pour tout $x \in M$.

b. $\rho(x, y) = \rho(y, x)$ quels que soient x et y de M.

c. $\rho(x, z) \leqslant \rho(x, y) + \rho(y, z)$ quels que soient x, y, z de M (axiome de triangle).

Une suite $x_1, x_2, \ldots$ de points d'un espace métrique M est dite *convergente* vers un point $x \in M$ si $\lim_{n\to\infty} \rho(x, x_n) = 0$.

12.22. Dans les chapitres précédents, on considérait en tant qu'exemples d'espaces métriques les ensembles sur la droite, dans le plan, dans l'espace habituel (euclidien) avec la métrique usuelle. Or, il importe de souligner que de divers ensembles de fonctions peuvent eux aussi être rendus espaces métriques en les munissant d'une métrique (i.e. d'une fonction $\rho(x, y)$) convenable.

Le choix de la métrique pour un espace fonctionnel dépend des exigences du problème considéré. Lorsqu'une distance est donnée, il est clair que deux éléments sont proches si la distance est petite. Dans la plupart des cas rencontrés dans l'analyse, on est forcé de procéder inversement: d'après les hypothèses d'un problème on voit

quels sont les éléments qu'il est naturel de considérer comme proches, ceci détermine la façon dont on introduit la métrique.

Par exemple, il est souvent naturel de considérer comme proches deux fonctions continues $x(t)$ et $y(t)$ $(a \leqslant t \leqslant b)$ pour lesquelles la quantité $\max\limits_{a \leqslant t \leqslant b} |x(t) - y(t)|$ est petite. On peut donc la choisir pour distance des fonctions $x(t)$ et $y(t)$; les axiomes a-c sont, évidemment, satisfaits et, par conséquent, tout ensemble M de fonctions continues sur l'intervalle $[a, b]$, avec pour distance

$$\rho(x, y) = \max_{a \leqslant t \leqslant b} |x(t) - y(t)|, \tag{1}$$

représente un espace métrique.

Dans certains cas (par exemple, dans le calcul des variations) où il s'agit des fonctions ayant leurs dérivées y compris la m-ième, il est naturel de considérer comme proches deux éléments $x(t)$ et $y(t)$ pour lesquels sont proches non seulement les valeurs des fonctions mais aussi celles des dérivées y compris la m-ième quel que soit t. Ceci conduit à la distance

$$\rho(x, y) = \max_{a \leqslant t \leqslant b} \{|x(t) - y(t)|, |x'(t) - y'(t)|, \ldots, |x^{(m)}(t) - y^{(m)}(t)|\}. \tag{2}$$

S'il y a un ensemble de fonctions $x(t)$ m fois continûment dérivables, alors, en le munissant de la métrique (2), on aboutit évidemment à un espace métrique.

Dans d'autres cas (par exemple, dans la théorie des équations intégrales), il est naturel de considérer les fonctions $x(t)$ et $y(t)$ comme proches si elles le sont au sens intégral, i.e. si la quantité

$$\int_a^b |x(t) - y(t)|\, dt$$

est petite.

Naturellement, on définit alors la distance par la formule

$$\rho(x, y) = \int_a^b |x(t) - y(t)|\, dt. \tag{3}$$

Il est évident que les axiomes de l'espace métrique sont satisfaits dans ce cas aussi.

Parfois on a besoin de définir la proximité des fonctions à l'aide de l'intégrale non pas de la différence de ces fonctions, mais d'une puissance, par exemple p-ième, de cette différence ; la distance correspondante peut être donnée par la formule

$$\rho(x, y) = \sqrt[p]{\int_a^b |x(t) - y(t)|^p\, dt}. \tag{4}$$

Pour $p \geqslant 1$, cette définition vérifie elle aussi les axiomes de l'espace métrique. Pour l'axiome c la vérification devient, pourtant, assez compliquée (à l'exception des cas simples $p = 1$ et $p = 2$); nous n'insistons pas là-dessus (cf. exercice 15).

Ainsi, la définition de l'espace métrique semble suffisamment souple pour satisfaire aux exigences les plus variées de l'analyse.

12.23. Espace des fonctions continues sur un espace métrique.

a. L'espace métrique de toutes les fonctions réelles continues sur l'intervalle $a \leqslant t \leqslant b$, avec la distance définie par la formule 12.22 (1), est noté $R^1[a, b]$ (comme dans 12.13t, où il figurait en tant qu'espace vectoriel).

b. Est-ce qu'il est possible de remplacer, dans cette définition, l'intervalle $[a, b]$ par n'importe quel espace métrique? Sur un espace métrique quelconque M, les fonctions continues ne sont pas nécessairement bornées, donc la formule 12.22 (1) de la distance ne convient plus. Or, on ne peut construire un espace fonctionnel qu'avec les fonctions continues et bornées, alors la formule 12.22 (1) conserve le sens, à condition que l'on y remplace max par sup. Définitivement, nous définissons $R^1(M)$ comme *espace de toutes les fonctions réelles continues et bornées sur un espace métrique* M, avec la distance

$$\rho(x, y) = \sup_{t \in M} |x(t) - y(t)| \tag{1}$$

entre les fonctions $x(t)$ et $y(t)$.

c. En remplaçant ici, à son tour, la droite réelle (domaine de valeurs des fonctions considérées) par un espace métrique quelconque P, on aboutit à l'espace $P^1(M)$ de toutes les fonctions continues et bornées sur un espace métrique M à valeurs dans un espace métrique P, avec la distance

$$\rho(x, y) = \sup_{t \in M} \rho_M\{x(t), y(t)\}. \tag{2}$$

entre les fonctions $x(t)$ et $y(t)$.

Dans les numéros suivants du présent paragraphe, nous étudions certaines notions générales de la théorie des espaces métriques relativement à l'espace $P^1(M)$ et à ses cas particuliers.

d. Il résulte de la définition (2) que la convergence d'une suite $x_n(t)$ vers la limite $x(t)$ dans l'espace $P^1(M)$ équivaut à la convergence *uniforme sur* M (5.93) de la suite des fonctions $x_n(t)$ vers la fonction limite $x(t)$.

e. Nous sommes convenus de dire qu'un ensemble E dans un espace métrique P est *partout dense par rapport à un ensemble* $F \subset P$ si tout point $x \in F$ ou bien appartient à E, ou bien en est un point limite (3.61). Si, de plus, $E \subset F$, on dit que E est partout dense dans F. Nous dirons qu'un espace métrique P est *séparable* s'il

existe un ensemble dénombrable $E \subset P$ *partout dense dans* P. Montrons que *l'espace* $R^s[a, b]$ *est séparable*.

Un ensemble E dénombrable et partout dense dans $R^s[a, b]$ peut être formé, par exemple de toutes les fonctions polygonales à sommets aux points (a, y_0), (x_1, y_1), . . ., (x_{n-1}, y_{n-1}), (b, y_n), où $x_j \in (a, b)$, les nombres x_j et y_j étant rationnels. La dénombrabilité de cet ensemble résulte de 2.35. Montrons que l'ensemble E est dense dans $R^s[a, b]$. Soient $f(x) \in R^s[a, b]$ une fonction quelconque et $\varepsilon > 0$. Trouvons un $\delta > 0$ tel que $|x' - x''| < \delta$ implique $|f(x') - f(x'')| < \varepsilon/5$. Soit $a = x_0 < x_1 < \ldots < x_n = b$ une partition de l'intervalle $[a, b]$ par les points rationnels x_j ($j = 1, \ldots, n-1$) et soit $\Delta x_j = x_{j+1} - x_j < \delta$. Supposons ensuite que les nombres rationnels $y_0, y_1, \ldots, y_n$ soient tels que $|y_j - f(x_j)| < \varepsilon/5$, $j = 0, 1, \ldots, n$. Considérons une fonction polygonale $y(x)$ aux sommets successifs (x_j, y_j). Alors $\rho(y, f) < \varepsilon$. En effet, pour tout $x \in [a, b]$, on peut trouver un x_j ($j = 1, \ldots, n-1$) tel que $|x - x_j| < \delta$. Alors $|f(x) - f(x_j)| < \varepsilon/5$ et $|f(x_{j\pm1}) - f(x_j)| < \varepsilon/5$. Il en résulte que

$$|y_j - y_{j\pm1}| \leqslant |y_j - f(x_j)| + |f(x_j) - f(x_{j\pm1})| + |f(x_{j\pm1}) - y_{j\pm1}| < \frac{3}{5}\varepsilon.$$

Donc, pour $x \in (x_{j-1}, x_{j+1})$, on a $|y(x) - y_j| < 3\varepsilon/5$. Définitivement,

$$|y(x) - f(x)| \leqslant |y(x) - y_j| + |y_j - f(x_j)| + |f(x_j) - f(x)| < \frac{3}{5}\varepsilon + \frac{1}{5}\varepsilon + \frac{1}{5}\varepsilon = \varepsilon,$$

$$\rho(y, f) = \max_{a \leqslant x \leqslant b} |y(x) - f(x)| < \varepsilon.$$

On peut prouver que l'espace $R^s[0, \infty)$ de toutes les fonctions bornées et continues sur la demi-droite $0 \leqslant x < \infty$ ne possède aucune partie dénombrable partout dense (cf. exercice 2).

f. Un espace métrique P est dit *complet* (3.71*d*) si le critère de Cauchy y est satisfait : toute suite de Cauchy $x_1, x_2, \ldots$ de P a une limite dans P.

Théorème. *L'espace* $P^s(M)$ *de toutes les fonctions continues et bornées sur un espace métrique* M, *à valeurs dans un espace métrique complet* P (*c*), *est un espace complet.*

Démonstration. Désignons par ρ la distance dans l'espace P et par

$$\rho(x, y) = \sup_t \rho_P\{x(t), y(t)\}$$

celle dans l'espace $P^s(M)$.

Soit $x_1(t), x_2(t), \ldots, x_n(t), \ldots$ une suite de Cauchy de fonctions éléments de l'espace $P^s(M)$: pour tout $\varepsilon > 0$, il existe un

numéro N tel que, pour $n \geqslant N$, $m \geqslant N$, on a l'inégalité

$$\rho(x_n, x_m) = \sup_t \rho_P\{x_n(t), x_m(t)\} \leqslant \varepsilon. \quad (3)$$

Il en résulte que toute suite $x_n(t_0) \in P$ qui s'obtient en fixant $t = t_0$ est une suite de Cauchy ; l'espace P étant complet, il existe la valeur

$$x(t_0) = \lim_{n\to\infty} x_n(t_0) \in P.$$

Lorsque $t = t_0$ parcourt tout le M, on a la fonction limite

$$x(t) = \lim_{n\to\infty} x_n(t).$$

Dans l'inégalité

$$\rho_P\{x_n(t), x_m(t)\} \leqslant \varepsilon,$$

qui a lieu pour tout t et n, $m \geqslant N$, passons à la limite lorsque $m \to \infty$ sans changer n ; nous aboutissons, en vertu de 5.12*b*, à

$$\rho_P\{x_n(t), x(t)\} \leqslant \varepsilon \quad (4)$$

pour tout t et pour $n \geqslant N$. Ceci veut dire que la suite des fonctions $x_n(t)$ converge vers la limite $x(t)$ uniformément sur M. D'après les théorèmes 5.94 et 5.95, la fonction $x(t)$ est bornée et continue, donc un élément de l'espace P^s (M). L'inégalité (4) peut être mise sous la forme

$$\rho(x_n, x) \leqslant \varepsilon,$$

ce qui veut dire que $x = x(t)$ est la limite de la suite des éléments x_n. Le théorème est démontré.

12.24. Théorème d'Arzelà. Un espace métrique P a été dit *compact* (3.91*a*) si chaque suite de ses points $x_1, x_2, \ldots$ contient une sous-suite $x_{m_1}, x_{m_2}, \ldots$ convergente, et *précompact* (3.93*a*) si chaque suite $x_1, x_2, \ldots$ contient une sous-suite de Cauchy $x_{m_1}, x_{m_2}, \ldots$.

a. Soient Q un espace métrique compact, P un espace métrique quelconque et P^s (Q) l'espace métrique de toutes les fonctions continues $x = x(t)$ définies sur Q et prenant leurs valeurs dans P métrisé d'après la formule 12.23 (2)

$$\rho(x, y) = \sup_t \rho_P\{x(t), y(t)\}.$$

L'espace P^s (Q) n'est en général pas compact. Dans quelles conditions un sous-ensemble $E \subset P^s$ (Q) est-il compact ?

Pour y répondre, introduisons les définitions suivantes :

Définition 1. Un ensemble E de fonctions $x(t) \in P^s$ (Q) est dit *à valeurs uniformément compactes* (*précompactes*) s'il existe un ensemble compact (précompact) $P_0 \subset P$ qui contient toutes les valeurs des fonctions $x(t)$ pour $t \in Q$, $x \in E$.

Définition 2. Un ensemble E de fonctions $x(t) \in P^a(Q)$ est dit *équicontinu* si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $\rho_Q(t', t'') < \delta$ implique

$$\rho_P\{x(t'), x(t'')\} < \varepsilon$$

quelle que soit la fonction $x(t) \in E$.

Théorème (Arzelà). *Pour qu'un ensemble $E \subset P^a(Q)$ soit précompact, il faut et il suffit qu'il soit à valeurs uniformément précompactes et équicontinu.*

Démonstration. Supposons que $E \subset P^a(Q)$ soit précompact. Alors, en vertu du critère de Hausdorff 3.93c, pour tout $\varepsilon > 0$ donné, il existe dans E un $\varepsilon/3$-réseau fini, i.e. un ensemble fini $x_1(t), \ldots, x_m(t)$ de fonctions tel que, pour toute fonction $x(t) \in E$, on peut trouver un numéro k, $1 \leqslant k \leqslant m$, pour lequel

$$\rho_P[x(t), x_k(t)] \leqslant \frac{\varepsilon}{3}. \tag{1}$$

Trouvons ensuite, pour le même ε, un $\delta > 0$ de façon à avoir les inégalités

$$\rho_P[x_k(t'), x_k(t'')] < \frac{\varepsilon}{3} \quad (k = 1, \ldots, m)$$

pour $\rho_Q(t', t'') < \delta$. Alors, pour les mêmes t' et t'', n'importe quelle $x(t) \in E$ et pour $x_k(t)$ correspondante, nous avons

$$\rho_P[x(t'), x(t'')] \leqslant \rho_P[x(t'), x_k(t')] + \rho_P[x_k(t'), x_k(t'')] +$$

$$+ \rho_P[x_k(t''), x(t'')] < 3 \cdot \frac{\varepsilon}{3} = \varepsilon,$$

c'est-à-dire que la famille E est équicontinue.

L'onsemble P_k de toutes les valeurs de la fonction $x_k(t)$ continue sur le compact Q est compact, quel que soit $k = 1, \ldots, m$ (5.16*a*). La réunion R des ensembles compacts $P_1, \ldots, P_m$ est évidemment compacte elle aussi. L'inégalité (1) montre que l'ensemble R sert d'$\varepsilon/3$-réseau à l'ensemble $P_0 \subset P$ de toutes les valeurs des fonctions $x(t) \in E$ sur Q. L'ensemble P_0 étant précompact en vertu de 3.95, E est à valeurs uniformément précompactes.

Nous voyons donc que les conditions du théorème d'Arzelà sont nécessaires pour la précompacité de E. Démontrons leur suffisance.

L'espace $P^a(Q)$ est isométriquement immergé dans l'espace $P(Q)$ de toutes les fonctions $x(t)$ bornées (continues ou non) sur Q muni de la distance

$$\rho(x, y) = \sup_t \rho_P\{x(t), y(t)\}.$$

En vertu du critère de Hausdorff (3.93c), le théorème sera démontré si l'on construit à partir de ses hypothèses, dans l'espace $P(Q)$

et pour tout $\varepsilon > 0$, un ε-réseau fini pour l'ensemble E. Supposons que $E \subset \mathrm{P}^s(\mathrm{Q})$ soit à valeurs uniformément précompactes et équicontinu.

Pour un $\varepsilon > 0$ donné, trouvons un $\delta > 0$ à partir de la condition d'équicontinuité de la famille E. Recouvrons ensuite le compact Q par un nombre fini de boules de diamètre δ. En rejetant les points superflus, on peut obtenir un recouvrement du compact Q par un nombre fini d'ensembles de diamètre $\leqslant\delta$ deux à deux disjoints. Désignons-les par $Q_1, \ldots, Q_m$. Soit ensuite $p_1, \ldots, p_k$ un $\varepsilon/2$-réseau fini dans P pour le précompact P_0 contenant toutes les valeurs des fonctions $x(t)$ pour $t \in \mathrm{Q}$, $x \in E$. Considérons l'ensemble G de toutes les fonctions $x(t) \in \mathrm{P}(\mathrm{Q})$ qui prennent sur $Q_1, \ldots, Q_m$ les valeurs constantes de la collection $p_1, \ldots, p_k$. Ces fonctions sont évidemment en nombre fini (au plus égal à k^m), et nous affirmons qu'elles forment un ε-réseau pour l'ensemble E. En effet, soit $x_0(t)$ une fonction quelconque de E. La variation de cette fonction sur l'ensemble Q_j de diamètre $\leqslant\delta$ ne dépasse pas $\varepsilon/2$, et il existe un point p_j de la collection $p_1, \ldots, p_m$ distant de toutes les valeurs de $x_0(t)$, pour $t \in Q_j$, de ε au plus. La fonction $x(t) \in \mathrm{P}(\mathrm{Q})$ qui prend sur chaque Q_j la valeur correspondante p_j appartient à G, et l'on a évidemment, dans l'espace $\mathrm{P}^s(\mathrm{Q})$:

$$\rho(x_0, x) = \sup_t \rho_P[x_0(t), x(t)] \leqslant \varepsilon.$$

Le théorème est démontré.

b. Si P est un espace complet, il en est de même de $\mathrm{P}^s(\mathrm{M})$ (12.23). Dans un espace complet, la fermeture de tout ensemble précompact est compacte (3.96*b*). Par conséquent, les parties compactes de $\mathrm{P}^s(\mathrm{M})$ sont caractérisées, dans le présent cas, comme suit : ce sont *les sous-ensembles fermés de* $\mathrm{P}^s(\mathrm{M})$ *à valeurs uniformément compactes et équicontinus.*

c. Dans le cas où P est l'espace euclidien n-dimensionnel R_n, la classe des ensembles précompacts est identique à celle des ensembles bornés (3.93*b* et 3.94). Le fait qu'une famille E de fonctions $x(t) \in R_n^s(\mathrm{M})$ est à valeurs uniformément précompactes veut donc dire, dans le cas en question, qu'il existe une constante B telle que $\sup|x(t)| \leqslant B$, quels que soient $t \in \mathrm{M}$ et $x(t) \in E$. Une telle famille de fonctions est dite *uniformément bornée*. Ainsi, dans le cas où $\mathrm{P} = R_n$, le théorème d'Arzelà prend la forme suivante :

Un ensemble E de fonctions $x(t) \in R_n^s(\mathrm{M})$ est précompact si, et seulement si, il est uniformément borné et équicontinu.

d. Pour les fonctions numériques sur un intervalle fermé de la droite numérique, on peut signaler une simple condition suffisante de la précompacité :

Théorème. *Si, pour un ensemble E de fonctions numériques $x(t)$ continues et dérivables sur un intervalle $a \leqslant t \leqslant b$, il existe*

deux constantes a_0, a_1 telles que, quel que soit $x(t) \in E$, on ait les inégalités

$$|x(t)| \leqslant a_0, \quad |x'(t)| \leqslant a_1,$$

alors l'ensemble E est précompact dans l'espace $R^s[a, b]$.

Démonstration. En appliquant la formule de Lagrange 7.44, nous avons l'inégalité

$$|x(t') - x(t'')| \leqslant \sup |x'(t)| \cdot |t'' - t'| \leqslant a_1 |t'' - t'|,$$

montrant que la famille E est équicontinue. L'application du théorème c achève la démonstration vu que ladite famille est uniformément bornée par hypothèse.

12.25. Espace des fonctions m fois continûment dérivables.

a. L'espace métrique de toutes les fonctions réelles $x(t)$ continues et m fois continûment dérivables sur un intervalle $a \leqslant t \leqslant b$ avec la métrique définie par la formule 12.22 (2)

$$\rho(x, y) = \max_{a \leqslant t \leqslant b} \{|x(t) - y(t)|, |x'(t) - y'(t)|, \ldots, |x^{(m)}(t) - y^{(m)}(t)|\}$$

est désigné par $D_m(a, b)$; en particulier, $D_0(a, b) = R^s(a, b)$. Dans l'espace $D_m(a, b)$, la convergence d'une suite $x_n(t)$ vers sa limite $x(t)$ signifie la convergence uniforme des $m + 1$ suites:

$$x_n(t) \to x(t), \quad x_n'(t) \to x'(t), \quad \ldots, \quad x_n^{(m)}(t) \to x^{(m)}(t).$$

b. Montrons que *l'espace $D_m(a, b)$ est complet.* Soit $x_1(t)$, $x_2(t)$, ... une suite de Cauchy de fonctions de l'espace $D_m(a, b)$. Il découle de l'inégalité

$$\max_{a \leqslant t \leqslant b} |x_n^{(k)}(t) - x_p^{(k)}(t)| \leqslant \rho(x_n, x_p)$$

que chacune des suites $\{x_n(t)\}$, $\{x_n'(t)\}$, ..., $\{x_n^{(m)}(t)\}$ est de Cauchy par rapport à la métrique de l'espace $R^s(a, b)$. L'espace $R^s(a, b)$ étant complet (12.23*f*), toute suite $x_n^{(k)}(t)$ converge uniformément, pour $n \to \infty$, vers une fonction continue $y_k(t)$ ($k = 0, 1, \ldots, m$). D'après le théorème 9.77 sur la dérivation d'une suite de fonctions, nous avons

$$y_1(t) = \lim_{n \to \infty} x_n'(t) = (\lim_{n \to \infty} x_n(t))' = y_0'(t),$$

$$y_2(t) = y_1'(t) = y_0''(t), \quad \ldots, \quad y_m(t) = y_0^{(m)}(t).$$

Donc, la fonction $y_0(t)$ appartient à l'espace $D_m(a, b)$. Toujours d'après la convergence uniforme de chaque suite $x_n^{(k)}(t)$ vers $y_k(t) = y_0^{(k)}(t)$, pour $n \to \infty$, la fonction $y_0(t)$ est la limite de la suite $x_n(t)$ par rapport à la métrique de l'espace $D_m(a, b)$, ce qu'il fallait démontrer.

c. Montrons que l'espace $D_m(a, b)$ est partout dense dans $R^s(a, b)$ (par rapport à la métrique de $R^s(a, b)$, bien entendu). La propriété d'être partout dense étant transitive (3.62), il suffit de montrer que $D_m(a, b)$ est partout dense par rapport à l'ensemble L de toutes les fonctions polygonales (nous avons vu dans 12.23*e* que L est partout dense dans $R^s(a, b)$). Toute fonction polygonale $y(x)$ peut être rendue « lisse » en la remplaçant, au voisinage de chaque angle, par une fonction $q(x)$ m fois continûment dérivable dont les valeurs des dérivées sont égales, aux points de contact avec les côtés du polygone, aux valeurs correspondantes des dérivées de la fonction $y(x)$ (fig. 12.1). Une telle fonction $q(x)$ peut être construite, par exemple, sous forme d'un polynôme de degré $2m$ (cf. 12.19*d*). En effectuant une homothétie de centre au sommet de l'angle et de rapport suffisamment petit, on peut arriver à ce que la distance entre $q(x)$ et $y(x)$ soit aussi petite que l'on veut.

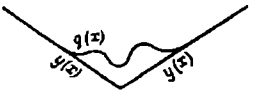


Fig. 12.1.

12.26. Espace des fonctions continues avec la métrique intégrale.

a. Désignons par $L^s(a, b)$ l'espace métrique de toutes les fonctions réelles $x(t)$ continues sur un intervalle fermé $[a, b]$ avec la métrique 12.22 (3)

$$\rho(x, y) = \int_a^b |x(t) - y(t)|\, dt. \tag{1}$$

L'espace formé des mêmes fonctions avec la métrique 12.22 (4)

$$\rho(x, y) = \sqrt[p]{\int_a^b |x(t) - y(t)|^p\, dt} \quad (p > 1) \tag{2}$$

sera désigné par $L_p^s(a, b)$.

Nous avons aussi considéré, sur l'espace des fonctions continues, la métrique

$$\rho(x, y) = \max_{a \leqslant t \leqslant b} |x(t) - y(t)|. \tag{3}$$

Les métriques (1)-(3) engendrent les types essentiellement différents de convergence. Ainsi, la suite des fonctions $x_n(t)$ montrées fig. 12.2 ne converge pas vers zéro dans l'espace $R^s(0, 1)$ tout en le faisant dans chacun des espaces $L_p^s(0, 1)$ parce que

$$\int_0^1 x_n^p(t)\, dt = \int_0^{1/n} x_n^p(t)\, dt \leqslant \frac{1}{n}.$$

La suite des fonctions $y_n(t)$ montrées fig. 12.3 converge vers zéro dans chaque espace $L_p^s(0, 1)$ avec $p < q$ et ne le fait pas dans $L_q^s(0, 1)$ parce que

$$\int_0^1 y_n^p(t)\,dt = \frac{1}{p+1}\cdot\frac{1}{n}\,n^{\frac{p}{q}} = \frac{1}{p+1}\,n^{\frac{p}{q}-1} \begin{cases} \to 0 & \text{pour } p<q, \\ = \frac{1}{p+1} & \text{pour } p=q. \end{cases}$$

b. *L'espace $L_p^s(a, b)$ n'est complet pour aucun $p \geqslant 1$.* Afin de le démontrer, considérons une suite de fonctions continues $y_\nu(x)$

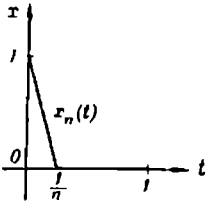


Fig. 12.2.

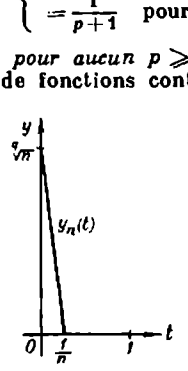


Fig. 12.3.

comprises entre 0 et 1 et convergeant, lorsque $\nu \to \infty$, vers 0 uniformément sur chaque intervalle $(a, c-\varepsilon)$ et vers 1 sur chaque intervalle $(c+\varepsilon, b)$ (c est un point fixe entre a et b). Une telle suite satisfait au critère de Cauchy dans l'espace $L_p^s(a, b)$. En effet, on a

$$\int_a^b |y_\nu(x) - y_\mu(x)|^p\,dx = \int_a^{c-\varepsilon} + \int_{c-\varepsilon}^{c+\varepsilon} + \int_{c+\varepsilon}^{b} \leqslant \varepsilon + 2\varepsilon + \varepsilon = 4\varepsilon$$

pour ν et μ assez grands. Montrons que *la suite $y_\nu(x)$ ne converge, selon la métrique de $L_p^s(a, b)$, vers aucune fonction continue.*

A cette fin, faisons la remarque suivante. Si une suite de fonctions $f_\nu(x)$ ($\nu = 1, 2, \ldots$) converge, par rapport à la métrique de $L_p^s(a, b)$, vers une fonction continue $f(x)$ sur un intervalle $\Delta = \{a \leqslant x \leqslant b\}$ et converge uniformément sur un intervalle $\delta = \{c \leqslant x \leqslant d\}$ intérieur à Δ vers une fonction $\varphi(x)$, alors l'identité $\varphi(x) \equiv f(x)$ a lieu dans l'intervalle δ. En effet, dans l'espace $L_p^s(c, d)$, nous avons les relations

$$\rho^p(f_\nu, f) = \int_c^d |f_\nu(x) - f(x)|^p\,dx \leqslant \int_a^b |f_\nu(x) - f(x)|^p\,dx \to 0$$

$$\rho^p(f_\nu, \varphi) = \int_c^d |f_\nu(x) - \varphi(x)|^p \, dx \leqslant \max_{x \in \delta} |f_\nu(x) - \varphi(x)|^p (d-c) \to 0.$$

En vertu de l'unicité de la limite (3.33*a*), on a bien $f(x) \equiv \varphi(x)$.

La supposition que la suite $y_1(x), y_2(x), \ldots, y_n(x), \ldots$ construite plus haut converge par rapport à la métrique de $L_p^s(a, b)$ vers une fonction continue $f(x)$ conduit, d'après ce qu'on a vu, aux égalités $f(x) = 0$ pour $a \leqslant x < c$ et $f(x) = 1$ pour $c < x \leqslant b$. Mais alors la fonction $f(x)$ ne peut être continue sur l'intervalle $a \leqslant x \leqslant b$ quelle que soit la valeur $f(c)$.

c. En vertu du théorème général 3.81, l'espace $L_p^s(a, b)$ admet le complété $\overline{L}_p^s(a, b)$. Il est naturel de poser la question suivante : est-il possible d'attribuer aux éléments de l'espace $\overline{L}_p^s(a, b)$, définis d'après le théorème 3.81*a* d'une façon abstraite, un sens concret en les interprétant, par exemple, comme fonctions? On répond par l'affirmative bien que le problème soit assez délicat (cf., par exemple, [16]).

§ 12.3. Espaces vectoriels normés

12.31. Nous tenons à munir un espace vectoriel R d'une métrique ; ce faisant, il est naturel d'admettre que la métrique et les opérations linéaires sont liées de sorte que la translation de deux points d'un même vecteur ne change pas la distance entre eux. Il est donc suffisant de définir la distance de tout point (vecteur) à un point fixe, au zéro par exemple. Il est suffisant, par conséquent, de faire correspondre à tout point $x \in R$ un nombre, la distance de x à 0 ; ce nombre est appelé *norme* du vecteur x.

Un espace vectoriel réel R s'appelle espace réel *normé* si, à tout vecteur $x \in R$, on fait correspondre un nombre $|x|$ (noté parfois $\|x\|$ ou même $|||x|||$) appelé *norme* du vecteur x et satisfaisant aux conditions :

a. $|x| > 0$ si $x \neq 0$, $|0| = 0$.

b. $|\alpha x| = |\alpha| \, |x|$ pour tout $x \in R$ et tout nombre réel α.

c. $|x + y| \leqslant |x| + |y|$ quels que soient x et y de R (axiome de triangle).

Par définition, posons $\rho(x, y) = |x - y|$. Il est aisé de prouver que les axiomes de l'espace métrique sont satisfaits (la preuve en est laissée au lecteur) ; il en résulte qu'un espace normé est métrique. Ainsi, dans un espace normé on peut mesurer les distances entre les vecteurs et se servir du passage à la limite. Un espace complet normé est dit *de Banach*. Un espace vectoriel qui n'est pas muni d'une norme (ou d'une métrique) sera dit *affine*.

12.32. a. L'espace $R^s(M)$ de toutes les fonctions réelles $x(t)$ continues et bornées sur un espace métrique M, que l'on a considéré

en tant qu'exemple d'un espace vectoriel (12.13*i*) et d'un espace métrique (12.23*b*), est en même temps une source importante d'espaces normés. Une norme dans R^s (M) est donnée par la formule

$$\| x \| = \sup_t | x(t) |. \qquad (1)$$

Nous désignons ici la norme d'une fonction $x(t)$ non pas par $| x |$ mais par $\| x \|$ pour mettre en relief la différence entre la norme de la fonction $x(t)$ comme élément de l'espace R^s (M) et sa valeur absolue dépendant du point t. Les axiomes 12.31 *a-c* de la norme sont dans ce cas presque évidents. En particulier, l'axiome de triangle se vérifie comme suit :

$$| x(t) + y(t) | \leqslant | x(t) | + | y(t) | \leqslant$$
$$\leqslant \sup_t | x(t) | + \sup_t | y(t) | = \| x \| + \| y \|;$$

en passant au maximum dans le premier membre, on obtient l'inégalité demandée $\| x + y \| \leqslant \| x \| + \| y \|$.

b. Généralisons maintenant l'exemple *a* en y remplaçant le corps R des nombres réels par un espace réel normé quelconque R. Ainsi, les éléments du nouvel espace R^s (M) seront les fonctions continues et bornées $x(t)$ définies sur l'espace métrique M à valeurs dans l'espace normé R.

Il est nécessaire de prouver que les opérations linéaires sur de telles fonctions conduisent aux fonctions de la même nature. Si les fonctions $x(t)$ et $y(t)$ à valeurs dans R sont bornées et que

$$X = \sup_t | x(t) |, \quad Y = \sup_t | y(t) |;$$

alors la fonction $z(t) = \alpha x(t) + \beta y(t)$ est aussi bornée pour n'importe quels α et β réels, puisque

$$| \alpha x(t) + \beta y(t) | \leqslant | \alpha | | x(t) | + | \beta | | y(t) | \leqslant$$
$$\leqslant | \alpha | X + | \beta | Y$$

quel que soit t.

Montrons que la fonction $z(t) = \alpha x(t) + \beta y(t)$ est continue pour tout $t = t_0$, de même que les fonctions $x(t)$ et $y(t)$. On peut supposer que $\alpha \neq 0$ et $\beta \neq 0$. Fixons un nombre $\varepsilon > 0$ et choisissons un $\delta > 0$ de façon que l'inégalité $\rho(t, t_0) < \delta$ implique

$$| x(t) - x(t_0) | < \frac{\varepsilon}{2\alpha}, \quad | y(t) - y(t_0) | < \frac{\varepsilon}{2\beta}.$$

On a alors

$$| z(t) - z(t_0) | \leqslant | \alpha | | x(t) - x(t_0) | + | \beta | | y(t) - y(t_0) | < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ce qui démontre la continuité de la fonction $z(t) = \alpha x(t) + \beta y(t)$ pour $t = t_0$.

Dans l'espace $R^s(M)$, on introduit la norme par la formule (1) où, bien entendu, $|x(t)|$ désigne la norme dans l'espace R.

Donc, l'espace $R^s(M)$ de toutes les fonctions $x(t)$ continues et bornées sur l'espace métrique M à valeurs dans l'espace normé R est construit.

Notons que l'espace $R^s(M)$ est complet dès que l'espace R l'est (12.23*f*).

12.33. Autres exemples d'espaces normés.

a. L'espace métrique $D_m(a, b)$ (12.25) peut être rendu normé en introduisant la norme

$$\|x\| = \max\{|x(t)|, |x'(t)|, \ldots, |x^{(m)}(t)|\}. \quad (1)$$

Les axiomes 12.31*a-c* de la norme sont vérifiés aisément.

b. Les espaces métriques $L_p^s(a, b)$ (12.26) peuvent être normés en introduisant les normes

$$\|x\|_p = \sqrt[p]{\int_a^b |x(t)|^p\, dt} \quad (p \geqslant 1). \quad (2)$$

Les axiomes 12.31*a-c* de la norme se vérifient aisément pour $p = 1$; dans le cas général ($p > 1$) la vérification de l'axiome de triangle demande une certaine patience (cf. exercice 15).

c. Les espaces normés $L_1^s(a, b)$ et $R^s(a, b)$ admettent des analogues intéressants de dimension finie. Soit R_n l'espace n-dimensionnel des vecteurs $x = (\xi_1, \ldots, \xi_n)$. Introduisons-y les normes suivantes :

$$|x|_1 = \sum_{k=1}^{n} |\xi_k|, \quad (3)$$

$$|x|_p = \sqrt[p]{\sum_{k=1}^{n} |\xi_k|^p} \quad (p > 1), \quad (4)$$

$$|x|_\infty = \max_{1 \leqslant k \leqslant n} |\xi_k|. \quad (5)$$

Il est évident que la norme euclidienne

$$|x| = \sqrt{\sum_{k=1}^{n} \xi_k^2} \quad (6)$$

que nous connaissons depuis 2.68 est un cas particulier de la norme $|x|_p$ correspondant à $p = 2$. Les normes (3) et (4) sont analogues aux normes intégrales dans les espaces $L_1^s(a, b)$ et $L_p^s(a, b)$. La norme (5) est analogue à la norme 12.32 (1) dans l'espace $R^s(a, b)$;

la désignation $|x|_\infty$ est justifiée par la relation limite

$$\max_{1\leqslant k\leqslant n}|\xi_k| = \lim_{p\to\infty}\sqrt[p]{\sum_{k=1}^{n}|\xi_k|^p}$$

(cf. exercice 9 du chapitre 4).

Dans le cas (6), nous avons déjà vérifié les axiomes de la norme dans 2.68. Dans les cas (3) et (5), cela se fait sans difficultés. Dans le cas (4), la vérification n'est plus si simple (cf. exercice 17).

Contrairement aux normes dans un espace fonctionnel, *du point de vue de la convergence qu'elles engendrent, les normes* (3)-(6) *sont équivalentes*: pour $m\to\infty$, la convergence d'une suite de vecteurs $x_m = (\xi_1^{(m)}, \ldots, \xi_n^{(m)})$ vers un vecteur $x = (\xi_1, \ldots, \xi_n)$ pour n'importe laquelle des normes (3)-(5) signifie la convergence de n suites numériques $\xi_1^{(m)}\to\xi_1, \ldots, \xi_n^{(m)}\to\xi_n$. L'étude des normes dans des espaces de dimension finie sera continuée dans 12.36.

d. En remplaçant dans les exemples précédents n par ∞ on aboutit à une série intéressante d'espaces de dimension infinie. Notamment, désignons par l_p $(1\leqslant p<\infty)$ l'ensemble de toutes les suites numériques $x=\{\xi_1, \xi_2, \ldots\}$ pour lesquelles $\sum_{k=1}^{\infty}|\xi_k|^p<\infty$. Soit $\|x\|_p = \sqrt[p]{\sum_{k=1}^{\infty}|\xi_k|^p}$. En se servant de l'inégalité de triangle pour la norme $|x|_p$ dans l'espace R_n et en posant $x=\{\xi_k\}\in l_p$, $y=\{\eta_k\}\in l_p$, on peut écrire

$$\sqrt[p]{\sum_{k=1}^{n}|\xi_k+\eta_k|^p}\leqslant\sqrt[p]{\sum_{k=1}^{n}|\xi_k|^p}+\sqrt[p]{\sum_{k=1}^{n}|\eta_k|^p}\leqslant$$

$$\leqslant\sqrt[p]{\sum_{k=1}^{\infty}|\xi_k|^p}+\sqrt[p]{\sum_{k=1}^{\infty}|\eta_k|^p}=\|x\|_p+\|y\|_p.$$

En passant à la limite pour $n\to\infty$ dans le premier membre on obtient la convergence de la série $\sum_{k=1}^{\infty}|\xi_k+\eta_k|^p$ et l'inéquation

$$\|x+y\|_p=\sqrt[p]{\sum_{k=1}^{\infty}|\xi_k+\eta_k|^p}\leqslant\|x\|_p+\|y\|_p.$$

Il est évident que $\alpha x = \{\alpha\xi_k\}$ appartient à l'ensemble l_p avec $x=\{\xi_k\}$, pour tout α réel, et que $\|\alpha x\|_p = |\alpha|\,\|x\|_p$. Ainsi *l'ensemble l_p, pour tout $p\geqslant 1$, est un espace vectoriel normé.* On peut montrer que l'espace l_p est complet, quel que soit $p\geqslant 1$ (exercice 18).

12.34. Toutes les définitions et tous les théorèmes concernant les espaces affines et les espaces métriques (sans structure d'espace vectoriel) sont naturellement valables pour les espaces vectoriels normés. Ainsi, dans un espace vectoriel normé on peut considérer les notions d'ensemble équilibré et d'ensemble convexe qui sont propres à la théorie des espaces affines.

a. Un ensemble E dans un espace vectoriel X est dit *équilibré* si, avec tout son point x, il contient le point $-x$.

b. Un ensemble E dans un espace vectoriel X est dit *convexe* si, avec ses deux points quelconques x, y, il contient tous les points

$$z = \alpha x + \beta y, \quad \alpha \geqslant 0, \quad \beta \geqslant 0, \quad \alpha + \beta = 1,$$

ou, géométriquement, tout le segment d'extrémités x et y.

c. Le théorème suivant utilise la structure d'espace vectoriel ainsi que la norme, de sorte que son champ d'action naturel est un espace vectoriel normé.

Théorème. *Une boule $S = \{x : |x| \leqslant \rho\}$ d'un espace vectoriel normé* R *est un ensemble équilibré convexe fermé.*

Démonstration. Si $|x| \leqslant \rho$, alors $|-x| = |x| \leqslant \rho$, ce qui démontre que la boule est équilibrée. Le fait qu'elle est fermée résulte de 3.51*b*. La convexité découle de l'axiome de triangle: si $|x| \leqslant \rho$, $|y| \leqslant \rho$, alors, pour $\alpha \geqslant 0$, $\beta \geqslant 0$, $\alpha + \beta = 1$, on a

$$|\alpha x + \beta y| \leqslant \alpha |x| + \beta |y| \leqslant (\alpha + \beta)\rho = \rho,$$

ce qu'il fallait démoutrer.

d. La propriété de convexité de la boule unité est si importante qu'elle peut remplacer l'axiome de triangle. Notamment, supposons que dans un espace vectoriel X on ait introduit une fonction numérique $|x|$ vérifiant les deux premiers axiomes de la norme et, au lieu du troisième, l'axiome suivant:

La boule $\{x \in X : |x| \leqslant 1\}$ est un ensemble convexe.

Montrons que cet axiome et les axiomes 12.31*a*, *b* impliquent l'inégalité de triangle 12.31*c*. Quels que soient deux vecteurs $x \neq 0$, $y \neq 0$, les vecteurs $\frac{x}{|x|}$ et $\frac{y}{|y|}$ appartiennent à la boule unité; en vertu du nouvel axiome, le vecteur $\frac{\alpha x}{|x|} + \frac{\beta y}{|y|}$ y appartient également quels que soient $\alpha \geqslant 0$, $\beta \geqslant 0$, $\alpha + \beta = 1$:

$$\left|\alpha \frac{x}{|x|} + \beta \frac{y}{|y|}\right| \leqslant 1.$$

Posons ici $\alpha = \frac{|x|}{|x|+|y|}$, $\beta = \frac{|y|}{|x|+|y|}$; en mettant $\frac{1}{|x|+|y|}$ en facteur, en le faisant sortir du symbole de la norme et en multipliant l'inégalité par $|x|+|y|$, on obtient

$$|x+y| \leqslant |x| + |y|,$$

ce qu'il nous fallait. Si l'un au moins des vecteurs x, y est nul, l'inégalité de triangle est immédiate.

12.35. Normes équivalentes.

a. Deux normes $|x|_1$ et $|x|_2$ dans un même espace vectoriel X sont dites *équivalentes* (ou *homéomorphes*) si les métriques qu'elles engendrent sont homéomorphes (3.34), autrement dit, si la convergence $x_n \to x$ selon l'une de ces normes est équivalente à la convergence $x_n \to x$ selon l'autre. Par conséquent, tout ensemble fermé (ouvert) dans X par rapport à l'une de ces normes est fermé (ouvert) par rapport à l'autre.

b. Voyons quelles propriétés géométriques des boules

$$S_1(\rho) = \{x \in X : |x|_1 \leqslant \rho\},$$

$$S_2(\rho) = \{x \in X : |x|_2 \leqslant \rho\}$$

correspondent à l'équivalence des normes $|x|_1$ et $|x|_2$. Chacune de ces boules, on l'a vu, est équilibrée, convexe et fermée par rapport à la norme correspondante.

Lemme. *Si les normes $|x|_1$ et $|x|_2$ sont équivalentes, alors il existe une constante $c_1 > 0$ telle que toute boule $S_1(\rho)$ contienne la boule $S_2(c_1\rho)$ et une constante $c_2 > 0$ telle que toute boule $S_2(\rho)$ contienne la boule $S_1(c_2\rho)$. Réciproquement, s'il existe deux constantes c_1 et c_2 ayant les propriétés mentionnées, alors les normes $|x|_1$ et $|x|_2$ sont équivalentes.*

Démonstration. Soient c_1 et c_2 des constantes possédant les propriétés en question. Soit ensuite $|x - x_n|_1 = \varepsilon_n \to 0$. La boule $S_2(\varepsilon_n/c_2)$ contenant $S_1(\varepsilon_n)$ contient l'élément $x - x_n$, de sorte que $|x - x_n|_2 \leqslant \varepsilon_n/c_2$. Ceci signifie que $|x - x_n|_2 \to 0$. D'une façon analogue, il résulte de $|x - x_n|_2 \to 0$ que $|x - x_n|_1 \to 0$.

Inversement, supposons que les normes $|x|_1$ et $|x|_2$ soient équivalentes, mais qu'il n'existe pas de constante c_1 cherchée. Alors, pour tout $n = 1, 2, \ldots$, nous trouvons deux boules $S_1(\rho_n)$, $S_2(\rho_n/n)$ dont la première ne contient pas la deuxième, c'est-à-dire qu'il existe un point x_n tel que $|x_n|_1 > \rho_n$, $|x_n|_2 < \rho_n/n$. Soit $y_n = x_n/\rho_n$; nous avons $|y_n|_1 > 1$, $|y_n|_2 < 1/n$, de sorte que la suite y_n tend vers 0 pour la deuxième norme et ne le fait pas pour la première. Ceci contredit la supposition d'équivalence des normes, d'où l'existence de la constante c_1. De même, il existe la constante c_2, ce qu'il fallait démontrer.

c. Conséquence. *Deux normes $|x|_1$ et $|x|_2$ sont équivalentes si, et seulement si, il existe deux constantes positives c_1 et c_2 telles que l'inégalité double*

$$c_1 |x|_1 \leqslant |x|_2 \leqslant \frac{|x|_1}{c_2}$$

soit satisfaite pour tout $x \in X$.

En effet, si l'inégalité en question est satisfaite, il résulte de $|x - x_n|_1 \to 0$ que $|x - x_n|_2 \leqslant \frac{1}{c_2}|x - x_n|_1 \to 0$ et inversement, de sorte que les normes $|x|_1$ et $|x|_2$ sont équivalentes. Supposons maintenant que les normes $|x|_1$ et $|x|_2$ soient équivalentes. Alors, d'après le lemme *b*, il existe une constante c_2 telle que toute boule $S_2(\rho)$ contient la boule $S_1(c_2\rho)$. Soit $|x|_1 = a$, c'est-à-dire que $x \in S_1(a)$. La boule $S_2(a/c_2)$ contient la boule $S_1(a)$, donc $|x|_2 \leqslant a/c_2 = |x|_1/c_2$. La deuxième inégalité s'établit de façon analogue.

d. A présent, nous sommes en mesure de décrire géométriquement n'importe quelle norme $|x|_2$ équivalente à une norme donnée $|x|_1$.

Théorème. *Supposons que, dans un espace normé* R *avec* $|x|_1$ *pour norme, on ait un ensemble équilibré convexe fermé* S *qui contient une boule* $S_1(\rho)$ *et est lui-même contenu dans une boule* $S_1(r)$. *Alors il existe une norme* $|x|_2$ *équivalente à la norme* $|x|_1$ *et telle que* $S_2(1) = S$.

Démonstration. Choisissons un vecteur quelconque $x \neq 0$ et considérons la demi-droite x/t, $0 < t < \infty$. Par hypothèse, les points de la demi-droite avec les t suffisamment grands appartiennent à l'ensemble S, et les points avec les t suffisamment petits n'y appartiennent pas. Posons $|x|_2 = \inf\left\{t : \frac{x}{t} \in S\right\}$ et $|0|_2 = 0$. Prouvons que la norme ainsi construite satisfait aux axiomes 12.31*a-c* et à la condition imposée $\{x : |x|_2 \leqslant 1\} = S$.

Pour $x \neq 0$ on a $0 < |x|_2 < \infty$, et le premier axiome est vérifié. Ensuite, pour $\alpha > 0$, on a

$$|\alpha x|_2 = \inf\left\{t : \frac{\alpha x}{t} \in S\right\} = \inf\left\{\alpha \frac{t}{\alpha} : \frac{\alpha x}{t} \in S\right\} =$$
$$= \alpha \inf\left\{\tau : \frac{x}{\tau} \in S\right\} = \alpha |x|_2.$$

L'ensemble S étant équilibré, il est immédiat que $|-x|_2 = |x|_2$; d'où $|\alpha x|_2 = |\alpha|\,|x|_2$ pour tout α réel.

Prouvons maintenant que $S = \{x : |x|_2 \leqslant 1\}$. Si $x \in S$, on a évidemment $|x|_2 = \inf\left\{t : \frac{x}{t} \in S\right\} \leqslant 1$. Notons ensuite que la convexité de S implique celle de l'ensemble S_x des points de la demi-droite x/t qui appartiennent à S; de la sorte, l'ensemble S_x contient *tous* les points x/t avec $t > \inf\left\{\tau : \frac{x}{\tau} \in S\right\}$; puisque S est fermé, le point x/t avec $t = \inf\left\{\tau : \frac{x}{\tau} \in S\right\}$ appartient lui aussi à S. Par conséquent, si $|x|_2 = \inf\left\{\tau : \frac{x}{\tau} \in S\right\} \leqslant 1$, alors $x = x/1$ appartient à S, ce qu'il nous fallait.

Vérifions l'inégalité de triangle pour la norme $|x|_2$. Elle résulte de 12.34*d*, car la boule $\{x: |x|_2 \leqslant 1\}$ est identique, on l'a vu, à l'ensemble S, ce dernier étant convexe. Il reste à démontrer que la norme $|x|_2$ est équivalente à $|x|_1$. Ceci découle du lemme *b*. En effet, l'inclusion $S_1(\rho) \subset S = S_2(1) \subset S_1(r)$ a pour conséquence l'inclusion $S_1(\rho p) \subset S_2(p) \subset S_1(rp)$ pour tout $p > 0$. L'hypothèse du lemme *b* étant donc satisfaite avec $c_1 = 1/r$ et $c_2 = \rho$; en appliquant le lemme on voit que les normes $|x|_2$ et $|x|_1$ sont équivalentes.

e. Normes dans les espaces de dimension finie. Montrons que *dans un espace vectoriel* R_n *de dimension finie toutes les normes sont équivalentes.*

La relation d'équivalence des normes est, évidemment, transitive. Il suffit donc de montrer que toute norme $|x|_1$ est équivalente à la norme euclidienne $|x|_2 = \sqrt{\sum_1^n \xi_i^2}$, où $\xi_1, \ldots, \xi_n$ sont les coordonnées du vecteur x dans une base $e_1, \ldots, e_n$.

Posons $c_1 = \sum_1^n |e_k|_1$; pour tout $x \in R_n$, on a l'inégalité

$$|x|_1 = \Big|\sum_1^n \xi_k e_k\Big|_1 \leqslant \sum_1^n |\xi_k|\,|e_k|_1 \leqslant |x|_2 \sum_1^n |e_k|_1 \leqslant c_1 |x|_2. \qquad (1)$$

Montrons qu'il existe également une constante c_2 telle que, pour tout $x \in R_n$, on a l'inégalité

$$|x|_1 \geqslant c_2 |x|_2. \qquad (2)$$

Supposons le contraire: il existe une suite de vecteurs x_m $(m = 1, 2, \ldots)$ telle que $|x_m|_1 < \frac{1}{m}|x_m|_2$. Posons $y_m = \frac{x_m}{|x_m|_2} = (\eta_1^{(m)}, \ldots, \eta_n^{(m)})$. Nous avons $|y_m|_2^2 = \sum_1^n (\eta_k^{(m)})^2 = 1$; d'où $|\eta_k^{(m)}| \leqslant 1$ quels que soient k et m. La boule euclidienne étant un ensemble compact (3.96*e*), la suite y_m $(m = 1, 2, \ldots)$ contient une sous-suite convergente; en rejetant les vecteurs superflus et en modifiant la numérotation on peut dire que la suite $y_m = (\eta_1^{(m)}, \ldots, \eta_n^{(m)})$ elle-même converge vers un vecteur $y = (\eta_1, \ldots, \eta_n)$. Alors on a, d'après 3.32*f*,

$$\eta_1 = \lim_{m\to\infty} \eta_1^{(m)}, \quad \ldots, \quad \eta_n = \lim_{m\to\infty} \eta_n^{(m)}. \qquad (3)$$

En passant à la limite pour $m \to \infty$ dans l'égalité $\sum_1^n (\eta_k^{(m)})^2 = 1$, on obtient $|y|_2^2 = \sum_1^n \eta_k^2 = 1$, de sorte que $y \neq 0$. En se basant

sur (1) on peut noter $|y-y_m|_1 \leqslant c_1 |y-y_m|_2 \to 0$, c'est-à-dire que $y_m \to y$ pour la norme $|x|_1$. Or $|y_m|_1 = \frac{|x_m|_1}{|x_m|_2} \to 0$, de sorte que $y_m \to 0$. Les relations obtenues contredisent l'unicité de la limite (3.33*a*). L'inégalité (2) est donc établie.

En appliquant maintenant la conséquence *c* notre affirmation se trouve démontrée.

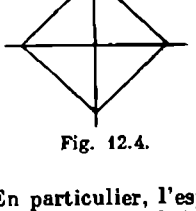


Fig. 12.4.

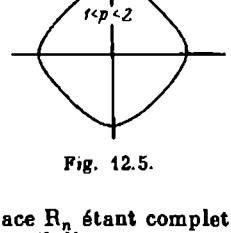


Fig. 12.5.

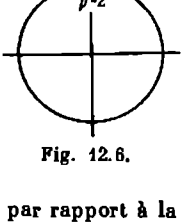


Fig. 12.6.

En particulier, l'espace R_n étant complet par rapport à la norme euclidienne $|x|_2$ (3.72*c*), il l'est par rapport à toute autre norme $|x|_1$.

f. En tant que conséquence nous obtenons:

Dans un espace R_n *de dimension finie, la convergence par rapport à n'importe quelle norme est équivalente à la convergence par rapport aux coordonnées.*

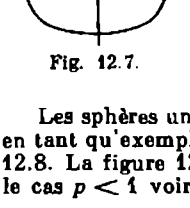


Fig. 12.7.

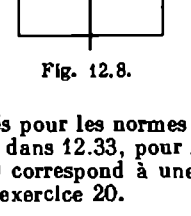


Fig. 12.8.

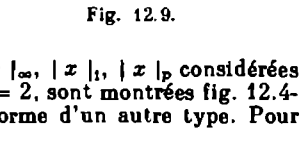


Fig. 12.9.

Les sphères unités pour les normes $|x|_\infty$, $|x|_1$, $|x|_p$ considérées en tant qu'exemples dans 12.33, pour $n = 2$, sont montrées fig. 12.4-12.8. La figure 12.9 correspond à une norme d'un autre type. Pour le cas $p < 1$ voir l'exercice 20.

12.36. Dans un espace euclidien *n*-dimensionnel, la boule $|x| \leqslant 1$ est compacte (3.96*e*). Dans tout espace normé *n*-dimensionnel, la boule $\|x\| \leqslant 1$ est aussi compacte, puisque toute norme est équivalente à la norme euclidienne d'après 12.35*e*. Est-ce qu'il existe des espaces normés de dimension infinie dans lesquels la boule unité $\|x\| \leqslant 1$ soit compacte? Il s'avère que non, et la

compacité de la boule est une propriété caractéristique des espaces de dimension finie.

a. Lemme. *Soit E un sous-espace fermé d'un espace vectoriel normé* R *et tel que* $E \neq$ R. *Il existe un vecteur* $y \in$ R *tel que* $|y| = 1$ *et* $|y - x| \geqslant 1/2$ *pour tous les* $x \in E$.

Démonstration. Choisissons un $y_0 \in$ R $- E$; soit $d = \inf |y_0 - x|$ pour $x \in E$. Si l'on avait $\inf |y_0 - x| = 0$, il existerait une suite $x_n \in E$ tendant vers y_0; E étant fermé, on aurait alors $y_0 = \lim x_n \in E$ en contradiction avec l'hypothèse. Donc $d > 0$. Trouvons un vecteur $x_0 \in E$ tel que $|y_0 - x_0| < 2d$. Posons $y = \frac{y_0 - x_0}{|y_0 - x_0|}$. Nous avons $|y| = 1$; de plus, $x_0 + x\,|y_0 - x_0| \in E$ pour tout $x \in E$ et

$$|y - x| = \left| \frac{y_0 - x_0}{|y_0 - x_0|} - x \right| = \left| \frac{y_0 - x_0 - x\,|y_0 - x_0|}{|y_0 - x_0|} \right| \geqslant \frac{d}{2d} = \frac{1}{2},$$

ce qu'il fallait démontrer.

b. Théorème (F. Riesz). *La boule unité d'un espace normé* R *de dimension infinie n'est pas un ensemble précompact.*

Démonstration. Nous allons construire, dans la boule unité de l'espace R, une suite $x_1, x_2, \ldots, x_n, \ldots$ des vecteurs éloignés l'un de l'autre d'une distance $> 1/2$. Comme une telle suite n'admet, évidemment, aucune sous-suite de Cauchy, la boule $S = \{x : |x| \leqslant 1\}$ n'est pas un ensemble précompact. Pour x_1 nous choisissons n'importe quel vecteur $x_1 \in S$, $|x_1| = 1$. Ses multiples λx_1 forment un sous-espace fermé $E_1 \subset$ R. D'après le lemme a, il existe un vecteur $x_2 \in S$, $|x_2| = 1$, tel que $|x_2 - x| > 1/2$ pour tout $x \in E_1$; en particulier $|x_2 - x_1| > 1/2$. Les combinaisons linéaires $\lambda_1 x_1 + \lambda_2 x_2$ forment un sous-espace fermé $E_2 \subset$ R. D'après le lemme a, il existe un vecteur $x_3 \in S$, $|x_3| = 1$, tel que $|x_3 - x| > 1/2$ pour tout $x \in E_2$; en particulier, $|x_3 - x_1| > 1/2$, $|x_3 - x_2| > 1/2$. En continuant nous aboutissons à une suite $E_1 \subset E_2 \subset \ldots$ de sous-espaces de dimension finie dont chacun est une partie propre de R (vu que ce dernier est de dimension infinie) et à une suite $x_1, x_2, \ldots$ des vecteurs avec les distances relatives $|x_m - x_n| > 1/2$. On a déjà vu que cela résout le problème.

12.37. Séries de vecteurs d'un espace normé. Dans un espace métrique, on peut considérer les suites convergentes, mais la notion de série convergente n'a pas de sens. Dans un espace vectoriel normé, la notion de série convergente de vecteurs a un sens.

a. Soit une série d'éléments d'un espace normé R :

$$x_1 + x_2 + \ldots + x_n + \ldots \qquad (1)$$

La série (1) est dite *convergente dans* R si la suite $s_1 = x_1$, $s_2 = x_1 + x_2, \ldots$ de ses sommes partielles converge dans R; dans

ce cas, la limite $s = \lim_{n\to\infty} s_n$ des sommes partielles est, par définition, la somme de la série (1). Si la suite des sommes partielles s_n n'est pas convergente, la série (1) est dite *divergente dans* R et on ne lui attribue aucune somme.

Pour la convergence de la série (1) *il faut et, si l'espace* R *est complet, il suffit que le critère de Cauchy soit satisfait: pour tout* $\varepsilon > 0$ *il existe un numéro* N *tel que l'inégalité*

$$| s_n - s_m | = | x_{m+1} + \ldots + x_n | < \varepsilon \qquad (2)$$

ait lieu quels que soient $m > N$, $n > m$.

b. *Si la série numérique des normes des vecteurs* x_n *converge, alors, dans le cas d'un espace* R *complet, la série* (1) *converge elle aussi*, puisque

$$| x_{m+1} + \ldots + x_n | \leqslant | x_{m+1} | + \ldots + | x_n |,$$

et on peut appliquer le critère de Cauchy.

c. Critère de Weierstrass. *La série* (1) *converge si les majorations* $|x_n| \leqslant \alpha_n$ *des normes ont lieu pour tous les* n (à partir d'un numéro quelconque) *et que la série numérique* $\sum_1^\infty \alpha_n$ *converge.*

En effet, dans l'hypothèse formulée la série $\sum_1^\infty |x_n|$ converge avec la sério $\sum_1^\infty \alpha_n$ d'après le critère de comparaison.

d. Critère de Cauchy. *La série* (1) *converge si* $\overline{\lim} \sqrt[n]{|x_n|} < 1$ *et diverge si* $\overline{\lim} \sqrt[n]{|x_n|} > 1$.

La démonstration se fait de même que dans 6.14*b* pour une série numérique.

e. Critère d'Abel-Dirichlet. *La série*

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + \ldots, \qquad (3)$$

où $x_1, x_2, \ldots$ *sont des vecteurs de l'espace* R *et* $a_1, a_2, \ldots$ *des nombres réels, converge dans* R *si les nombres* a_n *tendent vers zéro en décroissant d'une façon monotone et* $s_n = x_1 + \ldots + x_n$ *sont majorés en norme par une constante fixe.*

La démonstration suit la même voie que celle de 6.47 en remplaçant les modules par les normes.

f. Exemple. Considérons les séries

$$\sum_0^\infty a_n \cos nt, \qquad (4)$$

$$\sum_1^\infty b_n \sin nt \qquad (5)$$

dans l'espace $R^s(a, b)$. Rappelons que $R^s(a, b)$ est un espace complet (12.23*f*) et que la convergence pour la norme de l'espace $R^s(a, b)$ est la convergence uniforme sur l intervalle $[a, b]$. Les normes des fonctions $\cos nt$ et $\sin nt$ dans l'espace $R^s(a, b)$ sont au plus égales à l'unité. Donc, si $\sum |a_n| < \infty$ ou $\sum |b_n| < \infty$, alors la série (4) ou (5) respectivement converge dans l'espace $R^s(a, b)$ (d'après le critère de Weierstrass), i.e. converge uniformément sur $[a, b]$ quels que soient a et b.

En cas de divergence de la série formée par les a_n ou par les b_n et à condition que $a_n \searrow 0$ (ou $b_n \searrow 0$), on peut se servir du critère d'Abel-Dirichlet. Pour la somme des sinus ou des cosinus, nous avions (6.47 (9)) les majorations

$$\left| \sum_{m=0}^{n} \begin{matrix} \cos mt \\ \sin mt \end{matrix} \right| \leqslant \sqrt{\frac{2}{1-\cos t}}. \tag{6}$$

Si t varie dans l'intervalle $[\varepsilon, 2\pi - \varepsilon]$, où $\varepsilon > 0$, alors le second membre de l'inégalité (6) est borné, et dans le premier on peut passer au maximum :

$$\max \left| \sum_{0}^{n} \begin{matrix} \cos mt \\ \sin mt \end{matrix} \right| = \left\| \sum_{0}^{n} \begin{matrix} \cos mt \\ \sin mt \end{matrix} \right\| \leqslant \sqrt{\frac{2}{1-\cos \varepsilon}}.$$

Ceci garantit l'applicabilité du critère d'Abel-Dirichlet dans l'espace $R^s(\varepsilon, 2\pi - \varepsilon)$. Ainsi, les séries (4) et (5), dans l'hypothèse $a_n \searrow 0$, $b_n \searrow 0$, convergent uniformément sur tout intervalle $[\varepsilon, 2\pi - \varepsilon]$. Ces séries peuvent ne pas converger uniformément sur l'intervalle $[0, 2\pi]$ (bien que la série des sinus converge en tout son point !). Plus bas (14.47) nous verrons que

$$\sum_{1}^{\infty} \frac{1}{n} \sin nt = \begin{cases} 0 & \text{pour } t = 0 \text{ et } t = 2\pi, \\ \dfrac{\pi - t}{2} & \text{pour } 0 < t < 2\pi, \end{cases} \tag{7}$$

$$\sum_{1}^{\infty} \frac{1}{n} \cos nt = -\ln 2 \left| \sin \frac{t}{2} \right| \quad (0 < t < 2\pi). \tag{8}$$

Si la série (7) convergeait uniformément sur $[0, 2\pi]$, i.e. par rapport à la norme de l'espace $R^s(0, 2\pi)$, sa somme $s(t)$ appartiendrait à cet espace, donc serait une fonction continue sur $[0, 2\pi]$. Or, on voit de (7) que la fonction $s(t)$ est discontinue aux points 0 et 2π, donc la convergence uniforme sur l'intervalle $[0, 2\pi]$ de la série (7) n'a pas lieu.

La somme de la série (8) n'est pas bornée dans $[0, 2\pi]$, donc celle-ci n est pas non plus uniformément convergente sur $[0, 2\pi]$.

Aucune des séries (7) et (8) n'est uniformément convergente dans l'intervalle ouvert $(0, 2\pi)$.

12.38. Complété d'un espace normé. Comme chaque espace métrique, un espace normé R peut être complet ou non. Dans le dernier cas, l'espace R peut être complété en l'incluant dans un espace métrique complet plus vaste $\overline{R}$ (§ 3.8). De plus, le complété d'un espace normé est un espace non seulement métrique, mais aussi normé: nous introduirons dans le complété les opérations linéaires et vérifierons les axiomes de l espace normé.

Tout élément X du complété d'un espace métrique R était défini comme symbole correspondant à une classe de suites de Cauchy cofinales de l'espace R. Soit maintenant R un espace normé. Alors, en additionnant terme à terme deux suites de Cauchy $x_1, x_2, \ldots, x_n, \ldots$ et $y_1, y_2, \ldots, y_n, \ldots$, nous obtenons une suite

$$x_1 + y_1,\ x_2 + y_2,\ \ldots,\ x_n + y_n,\ \ldots,$$

qui est encore de Cauchy car

$$\| (x_n + y_n) - (x_m + y_m) \| \leqslant \| x_n - x_m \| + \| y_n - y_m \|.$$

Si l'on remplace ici la suite $\{x_n\}$ par une suite cofinale $\{x'_n\}$ et la suite $\{y_n\}$ par une suite cofinale $\{y'_n\}$, on aboutit à la suite des sommes $\{x'_n + y'_n\}$ qui est cofinale avec la suite $\{x_n + y_n\}$ car

$$\| (x'_n + y'_n) - (x_n + y_n) \| \leqslant \| x'_n - x_n \| + \| y'_n - y_n \|.$$

Ce fait permet de définir l'addition des éléments de l'espace $\overline{R}$. Choisissons, dans une classe X, une suite de Cauchy $\{x_n\}$ et, dans une classe Y, une suite de Cauchy $\{y_n\}$; appelons *somme* de X et Y la classe qui contient la suite de Cauchy $\{x_n + y_n\}$.

Les raisonnements qui précèdent confirment que la définition est correcte et que, en particulier, le résultat de l'addition ne dépend pas du choix des suites $\{x_n\}$ et $\{y_n\}$ dans les classes respectives.

D'une manière analogue, le *produit d'une classe X par un nombre* λ est défini comme suit: choisissons une suite de Cauchy $\{x_n\}$ dans la classe X et appelons λX la classe qui contient la suite de Cauchy $\{\lambda x_n\}$. Nous laissons au lecteur le soin de prouver que cette définition est correcte.

On vérifiera aisément les axiomes 12.11 de l'espace vectoriel; en vertu de la définition même, les opérations linéaires sur les classes se ramènent aux opérations correspondantes sur les éléments de l'espace initial. En particulier, la classe 0 est composée de toutes les suites convergentes vers 0 de l'espace R.

Il ne nous reste qu'à introduire une norme dans l'espace $\overline{R}$ et à vérifier les axiomes 12.31*a-c*. La norme d'une classe X est définie par la formule

$$\| X \| = \rho (X, 0),$$

où ρ signifie la distance dans l'espace métrique complété $\overline{R}$ (3.82 (1)). En d'autres termes, $\| X_n \| = \lim \rho (x_n, 0)$, où x_n est une suite de

Cauchy de la classe X. Si $\| X \| = 0$, alors $\lim \| x_n \| = 0$, de sorte que la suite $\{x_n\}$ est cofinale avec la suite $\{0, 0, \ldots\}$ définissant la classe 0; donc $X = 0$ et l'axiome 12.31*a* est vérifié. En fixant dans deux classes X, Y deux suites de Cauchy $\{x_n\}$, $\{y_n\}$ respectivement, nous choisissons dans la classe $X + Y$ la suite de Cauchy $\{x_n + y_n\}$. Etant donné $\| x_n + y_n \| \leqslant \| x_n \| + \| y_n \|$, il en découle

$$\| X + Y \| = \lim \| x_n + y_n \| \leqslant \lim \| x_n \| + \lim \| y_n \| = \\ = \| X \| + \| Y \|,$$

de sorte que l'axiome 12.31*c* est vérifié. D'une manière analogue, on a

$$\| \lambda X \| = \lim \| \lambda x_n \| = | \lambda | \lim \| x_n \| = | \lambda | \| X \|,$$

de sorte que l'axiome 12.31*b* est également vérifié. Ainsi notre affirmation est démontrée.

12.39. Espaces vectoriels complexes normés.

a. Dans 12.31-12.38 on considérait les espaces réels normés. Il n'est pourtant pas difficile d'introduire la notion d'un espace normé sur le corps des nombres complexes *). Notamment, un espace vectoriel complexe C est appelé *espace complexe normé* si à tout vecteur $x \in C$ on fait correspondre un nombre non négatif $| x |$, la *norme* du vecteur x, qui vérifie les conditions suivantes:

1) $| x | > 0$ si $x \neq 0$, $| 0 | = 0$;

2) $| \alpha x | = | \alpha | | x |$ pour tout $x \in C$ et pour tout α complexe;

3) $| x + y | \leqslant | x | + | y |$ quels que soient x et y de C (axiome de triangle).

Puisque, dans un espace complexe, la multiplication par tous les nombres complexes est admissible, tout espace complexe normé est en même temps un espace réel normé. Par conséquent, les propriétés des espaces réels normés peuvent être étendues directement ou sous une forme un peu modifiée aux espaces complexes normés. En particulier, un espace normé complexe, de même que réel, est un espace métrique avec la distance définie par la formule $\rho(x, y) = | x - y |$.

b. L'espace de toutes les fonctions $x(t)$ à valeurs complexes, bornées et continues sur un espace métrique M, avec pour norme

$$\| x \| = \sup_t | x(t) |,$$

*) Il est impossible d'étendre la définition au cas d'espace vectoriel sur un corps quelconque K car, pour les éléments α du corps quelconque K, la valeur absolue $| \alpha |$ n'est pas définie.

est un espace complexe normé que l'on désigne par $C^c(M)$. Cet espace est complet (12.23*f*).

c. L'espace des fonctions $x(t)$ complexes continues sur un intervalle $[a, b]$, muni de la norme

$$\| x \|_p = \sqrt[p]{\int_a^b | x(t) |^p \, dt},$$

est un espace complexe normé ; il est désigné par $CL_p^c(a, b)$ (ou tout simplement par $L_p^c(a, b)$, comme l'espace analogue des fonctions réelles, si aucune confusion n'est possible).

d. L'espace de toutes les fonctions $x(t)$ continues et bornées sur un espace métrique M, à valeurs dans un espace complexe normé C, muni de la norme

$$\| x \| = \sup_t | x(t) |$$

(où $| x(t) |$ est la norme de l'espace C), est un espace complexe normé ; il est désigné par $C^c(M)$. Il est complet si C est complet (12.23*f*).

e. Une petite modification suffit pour que les exemples d'espaces réels normés de dimension finie signalés dans 12.33*c* deviennent les exemples analogues d'espaces complexes normés de dimension finie : il n'y a qu'à remplacer le vecteur réel $x = (\xi_1, \ldots, \xi_n)$ par le vecteur complexe (i.e. considérer les coordonnées $\xi_1, \ldots, \xi_n$ comme nombres complexes) et à écrire, dans les formules 12.33 (6) et (4), $| \xi_k |^2$ et $| \xi_k |^p$ au lieu de ξ_k^2 et ξ_k^p. On obtient de même les analogues complexes des espaces l_p de dimension infinie (12.33*d*).

f. Un ensemble E dans un espace complexe normé C est dit *absolument convexe* si, avec ses deux points quelconques x, y, il contient tout point de la forme $\alpha x + \beta y$, où les nombres complexes α et β sont tels que $| \alpha | + | \beta | \leqslant 1$. Toute boule $\{x \in C : | x - x_0 | \leqslant \rho\}$ dans un espace complexe normé est un ensemble absolument convexe.

g. Les conditions d'équivalence des normes dans un espace réel normé (12.35*a-f*) restent valables pour les normes dans un espace complexe. En particulier, *dans un espace complexe de dimension finie n'importe quelles deux normes sont équivalentes, et la convergence par rapport à chaque norme est la convergence par rapport aux coordonnées.* Tous les espaces complexes normés de dimension finie sont complets, comme c'est le cas pour les espaces réels.

h. Une fois démontré pour les espaces réels le théorème de Riesz sur la non-compacité des boules dans un espace normé de dimension infinie (12.36*b*) est par là même établi pour les espaces complexes.

i. Toute la théorie de la convergence des séries de vecteurs d'un espace normé exposée dans 12.37 pour un espace réel s'étend, sans aucune modification, au cas d'un espace complexe.

Signalons ici un exemple spécifique. Soit une série de puissances

$$\sum_0^\infty a_k (z - z_0)^k,$$

où z et z_0 sont des nombres complexes et les coefficients a_k des éléments d'un espace complexe normé et complet C. Il s'avère que cette série *converge à l'intérieur du cercle de rayon*

$$r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\| a_n \|}}$$

centré au point z_0 et diverge à l'extérieur de ce cercle. On le démontre de même que la formule de Cauchy-Hadamard dans 6.62 en appliquant le critère de Cauchy 12.37*d*.

j. Le complété $\overline{C}$ d'un espace complexe normé C est construit de même que dans le cas réel (12.38) et représente un espace complexe normé et complet.

§ 12.4. Espaces hilbertiens

12.41. Dans un espace normé on peut mesurer les distances, mais non pas les angles, ce qui restreint les possibilités de l'interprétation géométrique. Dans un espace hilbertien on a, par définition, un produit scalaire de vecteurs par lequel on peut exprimer les longueurs des vecteurs ainsi que les angles qu'ils forment. Voici la définition rigoureuse : un espace vectoriel réel H s'appelle *espace hilbertien* si, pour deux vecteurs quelconques x et y de H, est défini un nombre réel (x, y) appelé *produit scalaire* des vecteurs x et y et satisfaisant aux conditions suivantes :

a. $(x, x) > 0$ si $x \neq 0$, $(0, 0) = 0$.

b. $(y, x) = (x, y)$ quels que soient x et y de H.

c. $(\alpha x, y) = \alpha (x, y)$ quels que soient x et y de H et un nombre réel α.

d. $(x + y, z) = (x, z) + (y, z)$ quels que soient x, y, z de H.

Les axiomes *b-d* ont pour conséquence la formule générale (par récurrence)

$$\Big(\sum_{j=1}^m \alpha_j x_j,\ \sum_{k=1}^n \beta_k y_k\Big) = \sum_{j=1}^m \sum_{k=1}^n \alpha_j \beta_k (x_j, y_k). \qquad (1)$$

L'axiomatique exposée se rapporte aux espaces hilbertiens réels ; celle des espaces hilbertiens complexes sera donnée plus bas (12.44).

12.42. Exemples.

a. L'espace euclidien n-dimensionnel R_n que nous avons introduit dans 2.68, avec le produit scalaire défini par la formule

$$(x, y) = \sum_1^n \xi_k \eta_k, \qquad (1)$$

où $x = \{\xi_1, \ldots, \xi_n\}$, $y = \{\eta_1, \ldots, \eta_n\}$, satisfait à toutes les conditions formulées ci-dessus.

b. L'espace n-dimensionnel R_n peut être muni d'un produit scalaire autre que (1). Il est aisé de caractériser tous les produits scalaires possibles dans R_n. Si (x, y) est un produit scalaire dans R_n et si $x = \sum \xi_k e_k$, $y = \sum \eta_k e_k$ sont les développements de deux vecteurs x et y dans une base $e_1, \ldots, e_n$ (2.75), alors, d'après la formule 12.41 (1), on a

$$(x, y) = \Big(\sum_k \xi_k e_k, \sum_j \eta_j e_j\Big) = \sum \xi_k \eta_j (e_k, e_j).$$

Ainsi, il suffit de connaître les valeurs du produit scalaire pour les vecteurs de la base (e_k, e_j) : le produit scalaire de deux vecteurs quelconques x, y sera déterminé d'une façon univoque d'après les nombres $\omega_{jk} = (e_j, e_k)$. Les nombres ω_{jk} doivent vérifier la condition de symétrie $\omega_{jk} = (e_j, e_k) = (e_k, e_j) = \omega_{kj}$ et l'inégalité

$$(x, x) = \sum \xi_j \xi_k \omega_{jk} > 0$$

pour tout $x \neq 0$; cela signifie que la matrice $\| \omega_{jk} \|$ doit être symétrique et définie positive. On démontre dans l'algèbre *) que les inégalités suivantes représentent une condition nécessaire et suffisante pour qu'une matrice symétrique $\| \omega_{jk} \|$ soit définie positive :

$$\omega_{11} > 0, \quad \begin{vmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{vmatrix} > 0, \quad \ldots, \quad \begin{vmatrix} \omega_{11} & \ldots & \omega_{1n} \\ \ldots & \ldots & \ldots \\ \omega_{n1} & \ldots & \omega_{nn} \end{vmatrix} > 0.$$

Inversement, toute matrice symétrique définie positive $\| \omega_{jk} \|$ définit, d'après la formule

$$(x \quad y) = \sum \xi_j \eta_k \omega_{jk},$$

un produit scalaire dans l'espace R_n, les axiomes 12.41*a-d* étant satisfaits. Après ce qu'on a dit, le lecteur réalisera facilement la démonstration.

c. Dans l'espace $R^1(a, b)$ de fonctions réelles continues, introduisons un produit scalaire, par exemple, d'après la formule suivante qui représente un analogue continu de la formule (1)

$$(x(t), y(t)) = \int_a^b x(t)\, y(t)\, dt. \tag{2}$$

Pour cette définition, la vérification des axiomes d'un espace hilbertien est immédiate, compte tenu des propriétés ordinaires de l'intégrale. (Il y a d'autres manières de munir l'espace $R^1(a, b)$ d'un produit scalaire.)

*) Cf. [14 ; 7.96].

d. Considérons l'espace vectoriel l_2 (12.33*d*) formé de toutes les suites numériques $x = \{\xi_1, \xi_2, \ldots\}$ telles que $\sum_{n=1}^{\infty} \xi_n^2 \infty 0$. Définissons le produit scalaire (x, y) de deux vecteurs $x = \{\xi_n\} \in l_2$ et $y = \{\eta_n\} \in l_2$ par la formule

$$(x, y) = \sum_{n=1}^{\infty} \xi_n \eta_n. \tag{3}$$

La convergence, même absolue, de la série du second membre découle de l'inégalité $|ab| \leqslant \frac{1}{2}(a^2 + b^2)$ qui est valable pour tout couple de nombres réels a et b. Les axiomes 12.41*a-d* sont ici immédiats.

Ainsi, l'espace l_2 devient espace hilbertien. La norme engendrée par le produit scalaire (3) coïncide avec la norme de l_2 introduite dans 12.33*d*.

12.43. Géométrie de l'espace hilbertien.

a. Dès 2.68, nous avons déduit l'inégalité de Cauchy-Bouniakovski

$$|(x, y)| \leqslant +\sqrt{(x, x)(y, y)} \tag{1}$$

pour deux vecteurs quelconques x et y d'un espace hilbertien H (puisque, en fait, nous n'avons utilisé que les axiomes d'un espace hilbertien).

Munissons l'espace hilbertien H de la norme

$$\|x\| = +\sqrt{(x, x)}. \tag{2}$$

Les axiomes 12.31 d'un espace normé se vérifient facilement : l'axiome 12.31*a* résulte de l'axiome 12.41*a*, 12.31*b* de 12.41*c*. En ce qui concerne l'axiome de triangle 12.31*c*, sa déduction des axiomes d'un espace hilbertien est, en fait, réalisée dans 2.68 en utilisant l'inégalité (1).

Ainsi, toutes les notions et propriétés liées à l'existence d'une norme sont valables pour un espace hilbertien. Cependant, celui-ci étant un cas très spécial de l'espace normé, nous avons le droit de s'attendre à ce que la norme dans l'espace hilbertion possède quelques propriétés spécifiques. L'une des propriétés de ce genre est donnée par le lemme suivant :

Lemme sur le parallélogramme : *Quels que soient deux vecteurs x et y d'un espace hilbertien* H, *on a l'égalité*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{3}$$

(« la somme des carrés des diagonales d'un parallélogramme est égale à la somme des carrés de ses côtés »).

La démonstration se réduit à une simple transformation :

$$\| x + y \|^2 + \| x - y \|^2 = (x + y,\ x + y) + (x - y,\ x - y) = \\ = 2\,(x,\ x) + 2\,(y,\ y) = 2\,\| x \|^2 + 2\,\| y \|^2.$$

On peut démontrer que si la norme dans un espace normé satisfait à la condition (3), elle est engendrée par un produit scalaire (exercice 4).

Quelle forme a la sphère $\{\| x \| = 1\}$ dans R_n dans le cas où la norme $\| x \|$ est obtenue à partir d'un produit scalaire $(x,\ y)$ d'après la formule $\| x \| = \sqrt{(x,\ x)}$ (cf. 2.68*a*) ?

Pour ce cas, nous avons

$$(x,\ x) = \Big(\sum_j \xi_j e_j,\ \sum_k \xi_k e_k\Big) = \sum_j \sum_k \xi_j \xi_k\,(e_j,\ e_k) = 1,$$

i.e. la sphère $\| x \| = 1$ est une surface à centre du deuxième degré ; comme elle est bornée, elle représente un ellipsoïde.

b. Soient $x_n \to x$ et $y_n \to y$ deux suites convergentes de vecteurs d'un espace hilbertien H. Montrons que

$$(x_n,\ y_n) \to (x,\ y).$$

En effet, nous avons

$$(x,\ y) - (x_n,\ y_n) = (x,\ y - y_n) + (x - x_n,\ y_n),$$

et, d'après l'inégalité (1)

$$|\,(x,\ y) - (x_n,\ y_n)\,| \leqslant \| x \|\,\| y - y_n \| + \\ + \| x - x_n \|\,\| y_n \|;$$

le deuxième membre tend vers 0 pour $n \to \infty$, car $\| y - y_n \| \to 0$, $\| x - x_n \| \to 0$, et la suite convergente y_n est bornée (3.33*b*).

c. Dans un espace hilbertien, on peut mesurer non seulement les longueurs des vecteurs (normes), mais aussi les angles qu'ils forment. L'*angle* de deux vecteurs non nuls x et y est défini par la formule

$$\cos \widehat{(x,\ y)} = \frac{(x,\ y)}{\| x \|\,\| y \|} ;$$

l'inégalité (1) assure l'existence de cet angle (dans l'intervalle $[0,\ \pi]$).

d. Deux vecteurs x et y d'un espace hilbertien H sont dits *orthogonaux* si $(x,\ y) = 0$. Lorsque $x \neq 0$ et $y \neq 0$, cette définition signifie que l'angle des vecteurs x et y vaut $\frac{\pi}{2}$. Le vecteur nul est orthogonal à tout vecteur.

Dans l'espace euclidien R_n avec le produit scalaire 12.42 (1), la condition d'orthogonalité de deux vecteurs

$$x = (\xi_1\ \ldots,\ \xi_n),\quad y = (\eta_1\ \ldots,\ \eta_n)$$

prend la forme

$$\sum_{k=1}^{n} \xi_k \eta_k = 0.$$

Dans l'espace fonctionnel $R^e(a, b)$ avec le produit scalaire 12.42 (2), la condition d'orthogonalité de deux vecteurs $x = x(t)$ et $y = y(t)$ a la forme

$$\int_a^b x(t)\, y(t)\, dt = 0.$$

e. Si un vecteur x est orthogonal aux vecteurs $y_1, \ldots, y_m$, il est orthogonal à toute leur combinaison linéaire $\alpha_1 y_1 + \ldots + \alpha_m y_m$. En effet,

$$(x,\ \alpha_1 y_1 + \ldots + \alpha_m y_m) = \alpha_1 (x,\ y_1) + \ldots$$
$$\ldots + \alpha_m (x,\ y_m) = 0.$$

Il en résulte que l'ensemble de *tous* les vecteurs orthogonaux à un vecteur x (ou bien à tout vecteur d'un ensemble fixe $X \in H$) forme un sous-espace dans H ; il s'appelle *supplémentaire orthogonal du vecteur* x (respectivement de l'ensemble X).

f. Théorème de Pythagore et sa généralisation. Supposons que deux vecteurs x et y soient orthogonaux ; alors, par analogie avec la géométrie élémentaire, on peut appeler le vecteur $x + y$ hypoténuse du triangle rectangle construit sur les vecteurs x et y. En formant le produit scalaire du vecteur $x + y$ par lui-même et en se servant de l'orthogonalité de x et y, on obtient

$$\| x + y \|^2 = (x + y,\ x + y) = (x,\ x) + 2\,(x,\ y) + (y,\ y) =$$
$$= \| x \|^2 + \| y \|^2.$$

Nous avons démontré, pour un espace hilbertien général, le théorème de Pythagore : *le carré de l'hypoténuse est égal à la somme des carrés des côtés*. Il est facile de généraliser ce théorème au cas d'un nombre quelconque de termes. A savoir, soient $x_1, \ldots, x_k$ des vecteurs deux à deux orthogonaux et $y = x_1 + \ldots + x_k$; alors on a

$$\| y \|^2 = (x_1 + \ldots + x_k,\ x_1 + \ldots + x_k) =$$
$$= \| x_1 \|^2 + \ldots + \| x_k \|^2.$$

g. Orthogonalisation. Pour obtenir un système de vecteurs orthogonaux, on utilise souvent l'orthogonalisation d'un système non orthogonal donné. Exposons-en la méthode. Soit un système de vecteurs $x_1, x_2, \ldots, x_n, \ldots$ d'un espace hilbertien H

dans lequel tout sous-système fini $x_1, \ldots, x_n$ est linéairement indépendant. *En se servant des formules*

$$\left.\begin{aligned} y_1 &= x_1, \\ y_2 &= a_{21}x_1 + x_2, \\ y_3 &= a_{31}x_1 + a_{32}x_2 + x_3, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{n,\,n-1}x_{n-1} + x_n, \end{aligned}\right\} \tag{4}$$

avec les coefficients a_{jk} convenablement choisis, on peut obtenir un système $y_1, \ldots, y_n, \ldots$ de vecteurs non nuls et deux à deux orthogonaux. Les formules (4) avec les coefficients a_{jk} dûment choisis sont appelées *formules d'orthogonalisation.*

L'existence d'une solution de ce système vérifiant les conditions demandées d'orthogonalité se démontre facilement par récurrence. En effet, supposons qu'on ait construit les vecteurs $y_1, \ldots, y_{n-1}$ non nuls et deux à deux orthogonaux qui vérifient les $n-1$ premières équations du système (4) et montrons qu'il est possible de trouver un vecteur y_n vérifiant la n-ième équation et orthogonal aux vecteurs $y_1, \ldots, y_{n-1}$. Cherchons le vecteur y_n comme combinaison linéaire des vecteurs $x_1, \ldots, x_n$ de la forme spéciale suivante :

$$y_n = b_{n1}y_1 + \ldots + b_{n,\,n-1}y_{n-1} + x_n, \tag{5}$$

où $y_1, \ldots, y_{n-1}$ sont les vecteurs déjà trouvés, $b_{n1}, \ldots, b_{n,\,n-1}$ les coefficients à déterminer. En multipliant scalairement l'équation (5) par y_k $(k < n)$ et en utilisant l'orthogonalité supposée de y_k à $y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_{n-1}$, nous obtenons

$$(y_n, y_k) = b_{nk}(y_k, y_k) + (x_n, y_k).$$

En égalant le second membre au zéro nous arrivons à une équation par rapport au coefficient b_{nk} qui est résoluble car $(y_k, y_k) \neq 0$ par supposition de récurrence. Lorsque tous les coefficients $b_{n1}, \ldots, b_{n,\,n-1}$ sont ainsi trouvés, l'égalité (5) détermine le vecteur y_n. Par construction, il sera orthogonal à chacun des vecteurs $y_1, \ldots, y_{n-1}$; il nous reste à démontrer que $y_n \neq 0$. Pour le faire, portons dans la n-ième équation (4) les expressions des $y_1, \ldots, y_{n-1}$ obtenues des $n-1$ premières équations ; nous aurons une expression linéaire de y_n par $x_1, \ldots, x_n$, où le coefficient devant x_n vaudra 1. Si y_n était nul, nous aurions une dépendance linéaire entre les $x_1, \ldots, x_{n-1}, x_n$, ce qui n'a pas lieu d'après la supposition de récurrence. Donc $y_n \neq 0$, et la méthode d'orthogonalisation est complètement établie.

On peut « perfectionner » le système orthogonal obtenu $y_1, \ldots, y_n, \ldots$ en divisant chaque vecteur y_n par sa longueur ; on

aura un système des vecteurs $e_n = \frac{y_n}{\|y_n\|}$ non seulement orthogonal, mais aussi normé, de sorte que chaque vecteur e_n aura 1 pour norme. En abrégé, un système orthogonal et normé de vecteurs est dit *orthonormé*.

h. Isomorphisme de deux espaces euclidiens n-dimensionnels. Conformément à la définition générale d'un isomorphisme de structures mathématiques (2.52), deux espaces hilbertiens H′ et H″ sont dits *isomorphes* s'ils sont isomorphes en tant qu'espaces vectoriels (12.14*i*) et si, de plus, les correspondances $x' \longleftrightarrow x''$, $y' \longleftrightarrow y''$ ($x', y' \in H', x'', y'' \in H''$) impliquent

$$(x', y') = (x'', y'').$$

Démontrons que *deux espaces hilbertiens quelconques d'une même dimension finie n sont isomorphes.*

Pour le faire, construisons dans un espace n-dimensionnel donné H_n une base orthonormée $e_1, \ldots, e_n$ en orthogonalisant par la méthode g un système linéairement indépendant quelconque de n vecteurs. Calculons le produit scalaire de deux vecteurs $x = \sum_1^n \xi_k e_k$ et $y = \sum_1^n \eta_m e_m$. Les vecteurs $e_1, \ldots, e_n$ étant orthonormés, nous avons

$$(x, y) = \left(\sum_1^n \xi_k e_k, \sum_1^n \eta_m e_m\right) = \sum_1^n \sum_1^n \xi_k \eta_m (e_k, e_m) = \sum_1^n \xi_k \eta_k. \qquad (6)$$

Ainsi, un espace hilbertien n-dimensionnel quelconque H_n peut être représenté comme espace de coordonnées (en faisant correspondre à tout vecteur $x = \sum_1^n \xi_k e_k$ l'ensemble $(\xi_1, \ldots, \xi_n)$ de ses coordonnées) avec le produit scalaire défini par (6). Or, ceci veut dire que *l'espace* H_n *est isomorphe à l'espace* R_n (12.42*a*). Deux espaces hilbertiens n-dimensionnels quelconques H_n' et H_n'' sont isomorphes, car chacun d'eux est isomorphe à un même espace R_n.

Le résultat démontré est fort important. En effet, même pour un espace hilbertien de dimension infinie, lorsqu'on opère dans un sous-espace de dimension finie, en particulier dans un espace bidimensionnel ou tridimensionnel, on peut se baser sur les propositions connues de la géométrie euclidienne ordinaire.

12.44. Comme l'analyste a souvent besoin de fonctions à valeurs complexes, il faut généraliser de façon convenable la notion d'espace hilbertien. Dans le cas où un espace vectoriel est complexe, les valeurs du produit scalaire que nous tenons à introduire peuvent

être complexes. Or, il n'est plus possible de conserver les conditions 12.41*a-c* parce que l'expression (ix, ix) doit être positive d'après *a*, tandis que, d'après *b* et *c*, on a

$$(ix, ix) = i(x, ix) = i(ix, x) = i^2(x, x) < 0.$$

Pour un espace complexe, nous acceptons la définition suivante.

Un espace vectoriel complexe (i.e. un espace vectoriel avec la multiplication par les nombres complexes) s'appelle *espace hilbertien* si, pour deux vecteurs quelconques x et y de H, est défini un nombre complexe (x, y) appelé produit scalaire de x et y et satisfaisant aux conditions suivantes:

a. $(x, x) > 0$ si $x \neq 0$; $(0, 0) = 0$.

b. $(y, x) = \overline{(x, y)}$ (nombre complexe conjugué) quels que soient x et y de H.

c. $(\alpha x, y) = \alpha (x, y)$ quels que soient x et y de H et α complexe.

d. $(x + y, z) = (x, z) + (y, z)$ quels que soient x, y, z de H.

Il résulte de *b* et *c*:

e. $(x, \alpha y) = \overline{(\alpha y, x)} = \bar{\alpha}\,\overline{(y, x)} = \bar{\alpha}(x, y)$.

A partir de *b-e*. on trouve facilement la formule générale

$$\Big(\sum_{j=1}^{m} \alpha_j x_j,\ \sum_{k=1}^{n} \beta_k y_k\Big) = \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \bar{\beta}_k (x_j, y_k). \qquad (1)$$

12.45. Exemples.

a. Le plus simple exemple d'un espace hilbertien complexe est fourni par l'espace complexe n-dimensionnel C_n. Il est formé des ensembles ordonnés de n nombres complexes $x = (\xi_1, \ldots, \xi_n)$, avec les opérations linéaires ordinaires (par coordonnées) et le produit scalaire défini comme suit: si $x = (\xi_1, \ldots, \xi_n)$, $y = (\eta_1, \ldots, \eta_n)$, alors $(x, y) = \xi_1\bar{\eta}_1 + \ldots + \xi_n\bar{\eta}_n$, où $\bar{\eta}_k$ est le nombre complexe conjugué de η_k. Les axiomes 12.44*a-d* sont immédiats.

On peut aussi munir l'espace C_n d'autres produits scaleires [14; § 9.1].

b. Un autre exemple d'espace hilbertien complexe est donné par l'espace $C^0[a, b]$ des fonctions $x(t)$ à valeurs complexes, continues sur un intervalle $a \leqslant t \leqslant b$, avec le produit scalaire défini par la formule

$$(x(t), y(t)) = \int_a^b x(t)\,\overline{y(t)}\,dt.$$

Les propriétés 12.44*a-d* résultent facilement des propriétés ordinaires de l'intégrale.

c. L'espace réel l_2 (12.42*d*) a pour analogue complexe l'espace de toutes les suites numériques complexes $x = \{\xi_n\}$ telles que

$\sum_{n=1}^{\infty} |\xi_n|^2 < \infty$. Ici le produit scalaire est donné par la formule

$$(x, y) = (\{\xi_n\}, \{\eta_n\}) = \sum_{n=1}^{\infty} \xi_n \bar{\eta}_n.$$

Les axiomes 12.44 se vérifient sans difficultés.

12.46.a. Soit H un espace hilbertien complexe. Posons, comme dans le cas réel,

$$\|x\| = +\sqrt{(x, x)}. \tag{1}$$

Démontrons l'inégalité de Cauchy-Bouniakovski-Schwarz

$$|(x, y)| \leqslant \|x\| \, \|y\|. \tag{2}$$

Pour tout α complexe, on a l'inégalité

$$(\alpha x - y, \alpha x - y) \geqslant 0.$$

En effectuant les opérations dans le premier membre, on obtient

$$\alpha\bar{\alpha}(x, x) - \alpha(x, y) - \bar{\alpha}\overline{(x, y)} + (y, y) \geqslant 0.$$

Posons $\alpha = te^{-i \arg (x, y)}$ (t réel); alors $\alpha(x, y) = t|(x, y)|$, et l'inégalité prend la forme

$$t^2(x, x) - 2t|(x, y)| + (y, y) \geqslant 0.$$

Comme le trinôme réel à gauche ne peut avoir de racines réelles distinctes (sinon il changerait de signe), ses coefficients satisfont à l'inégalité $|(x, y)|^2 \leqslant (x, x)(y, y)$, ce qu'il fallait démontrer.

b. Tout comme dans le cas réel, l'inégalité (2) implique l'inégalité de triangle

$$\|x + y\| \leqslant \|x\| + \|y\|$$

pour la norme (1).

c. Toujours comme dans le cas réel, deux vecteurs x, y d'un espace hilbertien complexe H sont dits *orthogonaux* si $(x, y) = 0$. Un système $x_1, x_2, \ldots, x_n, \ldots$ de vecteurs dont toute partie finie est linéairement indépendante peut être *orthogonalisé* comme dans 12.43*g*, c'est-à-dire qu'on peut construire d'après les formules 12.43*g* (4) un système de vecteurs non nuls deux à deux orthogonaux. En particulier, *tout espace hilbertien complexe n-dimensionnel* H_n *possède une base orthonormée* $e_1, \ldots, e_n$. Le produit scalaire des vecteurs $x = \sum_1^n \xi_k e_k$ et $y = \sum_1^n \eta_m e_m$ s'obtient selon la formule

$$(x, y) = \left(\sum_1^n \xi_k e_k, \sum_1^n \eta_m e_m\right) = \sum_1^n \xi_k \bar{\eta}_m (e_k, e_m) = \sum_1^n \xi_k \bar{\eta}_k. \tag{3}$$

En particulier, il résulte de la formule (3) que, de même que dans le cas réel, l'espace n-dimensionnel H_n est isomorphe à l'espace C_n (12.45*a*) et que, par conséquent, n'importe quels deux espaces hilbertiens complexes n-dimensionnels sont isomorphes.

12.47. Complété d'un espace hilbertien. Comme tout espace normé, un espace hilbertien H (réel ou complexe) peut être complet ou non. Ainsi, les espaces hilbertiens de dimension finie, réels ou complexes (12.42*a*, *b*, 12.45*a*), sont complets (12.35*e*, 12.39*g*). Les espaces de fonctions avec le produit scalaire intégral (12.42*c*, 12.45*b*) ne le sont pas (cf. 12.26*b*). Les espaces l_2, réel (12.42*d*) et complexe (12.45*c*), sont complets (exercice 18). Si un espace hilbertien H n'est pas complet, on peut le compléter en l'incluant dans un espace normé plus vaste comme nous l'avons fait dans 12.38. Montrons que *le complété d'un espace hilbertien est* non seulement un espace normé, mais aussi *un espace hilbertien.* Pour le faire, nous avons à définir dans le complété une opération de multiplication scalaire de façon que les axiomes 12.41*a-d* (dans le cas réel) ou 12.44*a-d* (dans le cas complexe) soient satisfaits.

Chaque élément X du complété d'un espace normé R était défini comme symbole correspondant à une classe de suites de Cauchy cofinales de l'espace R. Soient X et Y deux éléments quelconques du complété $\overline{H}$ d'un espace hilbertien H, $\{x_n\}$ et $\{y_n\}$ deux suites de Cauchy appartenant aux classes respectives. Montrons que les nombres (x_n, y_n) ont une limite pour $n \to \infty$. Nous avons

$$|(x_n, y_n) - (x_m, y_m)| = |(x_n - x_m, y_n) + (x_m, y_n - y_m)| \leqslant$$
$$\leqslant \|x_n - x_m\| \, \|y_n\| + \|x_m\| \, \|y_n - y_m\|.$$

Les suites de Cauchy $\{x_n\}$ et $\{y_n\}$ étant bornées (3.71*c*), la quantité obtenue tend vers zéro pour $m \to \infty$, $n \to \infty$, de sorte que la suite numérique (x_n, y_n) vérifie le critère de Cauchy. Il en résulte qu'elle possède une limite. Celle-ci ne dépend pas du choix de la suite $\{x_n\}$ dans la classe X et de la suite $\{y_n\}$ dans la classe Y; si $\{x'_n\}$ et $\{y'_n\}$ sont deux autres suites de ces classes, alors

$$|(x'_n, y'_n) - (x_n, y_n)| = |(x'_n - x_n, y'_n) - (x_n, y'_n - y_n)| \leqslant$$
$$\leqslant \|x'_n - x_n\| \, \|y'_n\| + \|x_n\| \, \|y'_n - y_n\| \to 0,$$

pour $n \to \infty$, de sorte que les suites numériques (x'_n, y'_n) et (x_n, y_n) ont une limite commune. Posons à présent

$$(X, Y) = \lim_{n \to \infty} (x_n, y_n).$$

Nous avons vu que le nombre (X, Y) est complètement déterminé par les classes X et Y sans dépendre du choix des suites $\{x_n\}$ et

$\{y_n\}$ dans ces classes. En particulier, le nombre $\sqrt{(X, X)} = \lim_{n\to\infty} \sqrt{(x_n, x_n)} = \lim_{n\to\infty} \| x_n \|$ coïncide avec la norme de la classe X dans l'espace normé $\bar{\mathrm{H}}$. L'axiome 12.41*a* est donc satisfait dans l'espace $\bar{\mathrm{H}}$. Les axiomes 12.41*b-d* (ou 12.44*b-d* dans le cas complexe) sont vérifiés en passant à la limite dans les axiomes respectifs pour l'espace H. Par exemple, dans le cas réel, on a

$$(Y, X) = \lim_{n\to\infty} (y_n, x_n) = \lim_{n\to\infty} (x_n, y_n) = (X, Y),$$

les autres axiomes étant vérifiés de façon analogue.

12.48. Espace préhilbertien.

a. Il arrive que, pour un espace vectoriel L que nous supposons réel pour fixer les idées, on peut introduire une fonction (x, y) d'une telle façon que les axiomes 12.41*b-d* se trouvent vérifiés et l'axiome 12.41*a* non : il existe des éléments $z \neq 0$ tels que $(z, z) = 0$. Un tel espace L s'appelle *espace préhilbertien*. Il s'avère possible de passer de l'espace L à un espace quotient L/E (12.14*i*) que l'on peut déjà considérer comme espace hilbertien.

b. Choisissons pour E l'ensemble de tous les éléments z tels que $(z, z) = 0$. Si $(z, z) = 0$ et $y \in \mathrm{L}$ est quelconque, alors, en vertu de l'inégalité de Cauchy-Bouniakovski dont la déduction ne se base que sur les axiomes 12.41*b-d*, on a

$$|(z, y)| \leqslant \sqrt{(z, z)}\sqrt{(y, y)} = 0, \tag{1}$$

de sorte que $(z, y) = 0$ pour tout $y \in \mathrm{L}$.

Montrons que E est un sous-espace dans L. Si $z_1 \in E$, $z_2 \in E$, alors, d'après (1),

$$(z_1 + z_2, z_1 + z_2) = (z_1, z_1) + 2(z_1, z_2) + (z_2, z_2) = 0,$$

donc $z_1 + z_2 \in E$. Il est aussi évident que $(z_1, z_1) = 0$ implique

$$(\alpha z_1, \alpha z_1) = \alpha^2 (z_1, z_1) = 0,$$

de sorte que $\alpha z_1 \in E$ dès que $z_1 \in E$. Par conséquent, l'ensemble E est un sous-espace dans L.

Formons l'espace quotient $\mathrm{H} = \mathrm{L}/E$ et munissons-le du produit scalaire

$$(X, Y) = (x, y), \tag{2}$$

où $x \in X$, $y \in Y$ sont arbitrairement choisis. Montrons avant tout que la définition donnée du produit scalaire ne dépend pas du choix des éléments x et y dans leurs classes. Soit $x_1 \sim x$, $y_1 \sim y$, de sorte que $x_1 = x + z$, $y_1 = y + u$, $z \in E$, $u \in E$. Alors, d'après (1), on a

$$(x_1, y_1) = (x, y) + (z, y) + (x, u) + (z, u) = (x, y),$$

ce qu'il nous fallait.

Vérifions les axiomes 12.41*a-d* pour l'espace H. Si $(X, X) = 0$, alors $(x, x) = 0$ pour tout $x \in X$, donc X coïncide avec la classe E qui sert de zéro à l'espace L/E (12.14*i*) de sorte que l'axiome 12.41*a* se trouve vérifié. En ce qui concerne les autres axiomes, ils résultent des axiomes respectifs de l'espace L et de la définition (2). Ainsi, l'espace $H = L/E$ est déjà un espace hilbertien.

c. On peut réaliser une construction parfaitement analogue pour un espace préhilbertien complexe L : si E est l'ensemble des $z \in L$ tels que $(z, z) = 0$, alors $L/E = H$ est un espace hilbertien complexe.

d. En guise d'exemple considérons l'espace vectoriel réel G (a, b) de toutes les fonctions continues par morceaux sur un intervalle $[a, b]$ avec pour produit scalaire

$$(x(t), y(t)) = \int_a^b x(t)\, y(t)\, dt.$$

Ici les axiomes 12.41*b-d* sont vérifiés et l'axiome 12.41*a* non, parce que, pour une fonction $z(t) \in G$ qui est nulle partout sauf en un nombre fini de points, nous avons

$$(z(t), z(t)) = \int_a^b z^2(t)\, dt = 0 \tag{3}$$

d'après 9.16*c*, bien que $z(t)$ ne soit pas le zéro de l'espace G. Par conséquent, G est un espace non pas hilbertien mais préhilbertien. On peut arriver à un espace hilbertien en passant de l'espace G à son espace quotient G/E, où E est l'ensemble de toutes les fonctions $z(t) \in G$ vérifiant l'égalité (3) : ce sont les fonctions qui ne diffèrent de zéro qu'en un nombre fini de points (9.16*d*). L'espace quotient G/E est formé des classes de fonctions $x(t) \in G$; deux fonctions appartiennent à une même classe si elles ne sont distinctes qu'en un nombre fini de points.

e. Le passage de l'espace préhilbertien complexe G $[a, b]$ de toutes les fonctions complexes continues par morceaux sur l'intervalle $[a, b]$, avec

$$(x(t), y(t)) = \int_a^b x(t)\, \overline{y(t)}\, dt$$

pour produit scalaire, à l'espace quotient hilbertien complexe G/E par le sous-espace E des fonctions complexes ne différant de zéro qu'en un nombre fini de points s'effectue d'une façon analogue.

Nous poursuivrons l'étude des espaces hilbertiens, du point de vue de leurs applications à l'analyse, dans le chapitre 14.

§ 12.5. Approximations dans l'espace des fonctions continues sur un compact

12.51. L'espace $R^*(Q)$ (resp. $C^*(Q)$) de toutes les fonctions réelles (resp. complexes) continues sur un compact Q est un espace vectoriel (12.13*i-k*) normé (12.32*a*, 12.39*d*) et complet (12.23 *f*). Nous considérerons de différentes familles linéaires $B(Q)$ de fonctions réelles (resp. complexes) continues sur le compact Q. *Quelles sont les conditions à imposer à la famille $B(Q)$ pour que sa fermeture par rapport à la convergence uniforme sur le compact Q, i.e. par rapport à la norme de l'espace $R^*(Q)$* (resp. *$C^*(Q)$*), *contienne toutes les fonctions continues sur Q*?

a. Nous dirons qu'une famille $B(Q)$ de fonctions *sépare deux points* z et y de l'ensemble Q s'il existe dans $B(Q)$ une fonction $\varphi(x)$ telle que $\varphi(z) \neq \varphi(y)$ (*fonction séparatrice pour les points z et y*). L'affirmation que $B(Q)$ ne sépare pas les points z et y signifie que $f(z) = f(y)$ pour toutes les fonctions $f(x) \in B(Q)$. Dans ce cas, vu que la dernière égalité est conservée en passant à la fermeture de la famille $B(Q)$ par rapport à la convergence uniforme, la fermeture de la famille $B(Q)$ ne peut contenir toutes les fonctions continues. Par exemple, elle ne contient pas la fonction $\rho(x, y)$ qui est nulle pour $x = y$ et non nulle pour $x = z$. Donc, si nous voulons que la fermeture d'une famille $B(Q)$ contienne toutes les fonctions continues sur le compact Q, nous avons à supposer qu'elle sépare n'importe quels deux points du compact Q.

b. Une famille linéaire $B(Q)$ de fonctions réelles sur l'ensemble Q est appelée *réseau linéaire* si, avec toute sa fonction $f(x)$, l'ensemble $B(Q)$ contient la fonction $|f(x)|$.

Quels que soient deux nombres réels α et β, on a

$$\max\{\alpha, \beta\} + \min\{\alpha, \beta\} = \alpha + \beta,$$
$$\max\{\alpha, \beta\} - \min\{\alpha, \beta\} = |\alpha - \beta|.$$

Par conséquent, quelles que soient deux fonctions réelles $f(x)$, $g(x)$, on a

$$\max\{f(x), g(x)\} + \min\{f(x), g(x)\} = f(x) + g(x),$$
$$\max\{f(x), g(x)\} - \min\{f(x), g(x)\} = |f(x) - g(x)|.$$

En résolvant ces équations par rapport à $\max\{f(x), g(x)\}$ et $\min\{f(x), g(x)\}$ nous voyons que, avec deux fonctions $f(x)$ et $g(x)$, un réseau linéaire contient les fonctions $\max\{f(x), g(x)\}$ et $\min\{f(x), g(x)\}$. Ensuite, en raisonnant par récurrence on conclut aisément que si un réseau linéaire contient des fonctions $f_1(x), \ldots, f_n(x)$, il contient aussi les fonctions $\max\{f_1(x), \ldots, f_n(x)\}$ et $\min\{f_1(x), \ldots, f_n(x)\}$.

c. Théorème. *Un réseau linéaire $B(Q)$ sur un compact Q est partout dense dans l'espace $R^*(Q)$ de toutes les fonctions continues sur Q s'il sépare n'importe quels deux points du compact et contient la fonction $e(x) \equiv 1$.*

Démonstration. Un réseau linéaire $B(Q)$ contenant 1 et séparant deux points z et y contient toute fonction qui prend aux points z et y n'importe quelles valeurs données; on peut trouver une telle fonction sous la forme $a\varphi(x) + b \cdot 1$, où $\varphi(x)$ est une fonction de $B(Q)$ qui sépare les points z et y et a, b des constantes.

Donnons-nous un $\varepsilon > 0$ et une fonction continue $f(x)$. Quels que soient deux points z et y (pas nécessairement distincts), on peut trouver d'après ce qu'on a dit plus haut une fonction $\varphi_{zy}(x) \in B(Q)$ pour laquelle $\varphi_{zy}(z) = f(z)$, $\varphi_{zy}(y) = f(y)$. Soit

$$U_{zy} = \{x \in Q: \varphi_{zy}(x) < f(x) + \varepsilon\}.$$

L'ensemble U_{zy} est ouvert et contient les points z et y. Fixons z; alors les ensembles ouverts U_{zy} considérés pour tous les $y \in Q$ forment un recouvrement du compact Q. D'après le lemme 3.97 on peut en extraire un recouvrement fini $U_{zy_1}, \ldots, U_{zy_m}$. Considérons la fonction

$$\varphi_z(x) = \min\{\varphi_{zy_1}(x), \ldots, \varphi_{zy_m}(x)\}$$

appartenant au réseau linéaire $B(Q)$. Vu que, pour tout $x \in Q$ et pour z fixe, l'une au moins des inégalités définissant les domaines U_{zy_k} est valable, nous avons $\varphi_z(x) \equiv \min_k \varphi_{zy_k} < f(x) + \varepsilon$ pour tout $x \in Q$. En même temps nous avons $\varphi_z(z) = \min_k \varphi_{zy_k}(z) = f(z)$.

Posons

$$V_z = \{x \in Q: \varphi_z(x) > f(x) - \varepsilon\}.$$

L'ensemble V_z est ouvert et contient le point z. Les ensembles V_z pour tous les $z \in Q$ forment un recouvrement du compact Q. D'après le lemme 3.97 on peut en extraire un recouvrement fini

$$V_{z_1}, \ldots, V_{z_n}.$$

Posons maintenant

$$\varphi(x) = \max\{\varphi_{z_1}(x), \ldots, \varphi_{z_n}(x)\};$$

cette fonction appartient elle aussi au réseau linéaire $B(Q)$, et l'on a par construction

$$\varphi(x) = \max_j \varphi_{z_j}(x) < f(x) + \varepsilon.$$

Or, en tout point $x \in Q$ l'une au moins des inégalités définissant le domaine V_{z_j} est valable, donc

$$\varphi(x) = \max \varphi_{z_j}(x) > f(x) - \varepsilon.$$

Définitivement, pour tout $x \in Q$ on a

$$f(x) - \varepsilon < \varphi(x) < f(x) + \varepsilon,$$

ce qui démontre le théorème.

(Le théorème tombe en défaut en rejetant l'hypothèse $e(x) \equiv 1 \in B(Q)$: étant donnés deux points z et y, le réseau linéaire de toutes les fonctions continues $f(x)$ vérifiant la condition $f(z) = 2f(y)$ n'est pas partout dense dans l'espace $R^e(Q)$.)

12.52. Théorème de Stone.

a. Conformément à la définition générale d'une algèbre (12.18*a*), une famille linéaire $B(Q)$ composée de fonctions (réelles) sur un compact Q s'appelle *algèbre* si, avec ses deux fonctions quelconques $f(x)$ et $g(x)$, la famille $B(Q)$ contient la fonction $f(x)\,g(x)$.

b. Lemme. *Une algèbre réelle $B(Q)$ qui contient l'unité et est fermée par rapport à la convergence uniforme représente un réseau linéaire.*

Démonstration. Démontrons que la fonction $|f(x)|$ appartient à l'algèbre $B(Q)$ dès que $f(x)$ le fait. Sans porter atteinte à la généralité on peut poser $\max_x |f(x)| = 1$.

Considérons la série de Taylor

$$(1-\xi)^{1/2} = 1 - \frac{1}{2}\xi + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\xi^2}{1\cdot 2} - \ldots$$

$$\ldots + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\ldots\left(\frac{1}{2}-n+1\right)}{1\cdot 2\cdot 3\ldots n}(-\xi)^n + \ldots$$

On a vu dans 9.52*d* (il faut y poser $\alpha = 1/2$) et dans 6.65 qu'elle converge uniformément pour $0 \leqslant \xi \leqslant 1$.

L'inégalité $0 \leqslant f^2(x) \leqslant 1$ étant vérifiée sur le compact Q, on a d'après ce qu'on a dit plus haut :

$$|f(x)| = \sqrt{1-(1-f^2(x))} =$$

$$= 1 - \frac{1}{2}(1-f^2(x)) - \frac{1}{8}(1-f^2(x))^2 + \ldots,$$

la série à droite étant uniformément convergente sur Q. Comme l'algèbre $B(Q)$ est fermée par rapport à la convergence uniforme, $|f(x)| \in B(Q)$, ce qu'il était à démontrer.

c. Théorème de Stone (pour une algèbre réelle). *Une algèbre $B(Q)$ formée de fonctions réelles, séparant n'importe quels deux points du compact Q et contenant l'unité est partout dense dans l'espace $R^e(Q)$.*

Démonstration. Désignons par $\overline{B(Q)}$ la fermeture de l'algèbre $B(Q)$ par rapport à la convergence uniforme. Evidem-

ment, la famille $\overline{B(Q)}$ est encore une algèbre : si $f_n(x) \to f(x)$ (uniformément sur Q) et que $g_n(x) \to g(x)$ (uniformément sur Q), alors $f_n(x)\, g_n(x) \to f(x)\, g(x)$ (uniformément sur Q) de sorte que $f(x)\, g(x) \in \overline{B(Q)}$ résulte de $f(x) \in \overline{B(Q)}$, $g(x) \in \overline{B(Q)}$.

L'algèbre $\overline{B(Q)}$ est un réseau linéaire (lemme *b*) et partout dense dans l'espace $R^e(Q)$ (théorème 12.51*c*). Comme l'algèbre $\overline{B(Q)}$ est fermée, $\overline{B(Q)} = R^e(Q)$, ce qu'il fallait démontrer.

12.53. a. On pourrait s'attendre à ce qu'une algèbre formée de fonctions à valeurs complexes, à condition qu'elle sépare n'importe quels deux points du compact Q et contienne l'unité, soit partout dense dans l'espace $C^e(Q)$ de toutes les fonctions complexes continues sur Q. Cependant, sous cette forme, le théorème s'avère faux (cf. exercice 5).

b. Tout de même, dans une hypothèse supplémentaire, le théorème de Stone s'étend aux algèbres de fonctions à valeurs complexes. Une algèbre complexe $B(Q)$ est dite *symétrique* si, avec toute sa fonction $\varphi(x) = u(x) + iv(x)$, elle contient la fonction conjuguée $\overline{\varphi}(x) = u(x) - iv(x)$.

Théorème de Stone (pour une algèbre complexe). *Une algèbre $B(Q)$ formée de fonctions à valeurs complexes, séparant n'importe quels deux points du compact Q, contenant l'unité et symétrique est partout dense dans l'espace $C^e(Q)$.*

Démonstration. Par hypothèse, l'algèbre $B(Q)$ contient, avec une fonction $\varphi(x) = u(x) + iv(x)$, les fonctions réelles $u(x) = \frac{1}{2}[\varphi(x) + \overline{\varphi}(x)]$ et $v(x) = \frac{1}{2i}[\varphi(x) - \overline{\varphi}(x)]$. Désignons par $B_R(Q)$ la sous-algèbre des fonctions réelles $h(x) \in B(Q)$. Cette sous-algèbre sépare n'importe quels deux points y et z du compact Q (si $\varphi(y) \neq \varphi(z)$, alors ou bien $u(y) \neq u(z)$ ou bien $v(y) \neq v(z)$) et contient l'unité. D'après le théorème de Stone 12.52*c*, on a $\overline{B_R(Q)} = R^e(Q)$, d'où $\overline{B(Q)} = C^e(Q)$.

12.54. Conséquences des théorèmes de Stone.

a. Supposons que le compact Q soit une partie fermée bornée de R_n et que l'algèbre $B(Q)$ soit formée de tous les polynômes réels $p(x_1, \ldots, x_n)$. Toutes les hypothèses du théorème de Stone 12.52*c* sont évidemment vérifiées. En l'appliquant on aboutit au théorème suivant :

Théorème (Weierstrass). *Toute fonction réelle $f(x)$ continue sur un ensemble fermé borné $Q \subset R_n$ est la limite uniforme sur Q d'une suite de polynômes en $x_1, \ldots, x_n$.*

b. Pour l'algèbre $B(Q)$ des polynômes $p(x_1, \ldots, x_n)$ à valeurs complexes, les hypothèses du théorème de Stone 12.53*b* sont véri-

fiées ; par conséquent, *toute fonction à valeurs complexes $f(x)$ continue sur un ensemble fermé borné $Q \subset R_n$ est la limite uniforme sur Q d'une suite de polynômes (à valeurs complexes) en $x_1, \ldots, x_n$.*

c. En particulier, *toute fonction (réelle ou complexe) continue sur un intervalle fermé $a \leqslant x \leqslant b$ est la limite uniforme d'une suite de polynômes (respectivement réels ou complexes) en x.*

d. Supposons à présent que le compact Q soit la circonférence $x^2 + y^2 = 1$ dans le plan des x, y. La position d'un point sur cette circonférence est déterminée par l'angle polaire φ. Choisissons pour algèbre $B(Q)$ l'ensemble des polynômes trigonométriques à coefficients réels

$$p(\varphi) = \sum_{k=0}^{n} (a_k \cos k\varphi + b_k \sin k\varphi). \tag{1}$$

Les formules de multiplication de fonctions trigonométriques (5.63) que l'on peut mettre sous la forme

$$2 \cos k\varphi \cos m\varphi = \cos (k - m)\,\varphi + \cos (k + m)\,\varphi,$$
$$2 \cos k\varphi \sin m\varphi = \sin (m - k)\,\varphi + \sin (m + k)\,\varphi,$$
$$2 \sin k\varphi \sin m\varphi = \cos (k - m)\,\varphi - \cos (k + m)\,\varphi$$

impliquent que l'ensemble des fonctions (1) contient, avec ses deux fonctions quelconques, leur produit, donc est effectivement une algèbre. N'importe quels deux points φ_1 et φ_2 sont séparés par une fonction de l'algèbre $B(Q)$, notamment par $\sin\varphi$ ou $\cos\varphi$. En appliquant le théorème de Stone 12.52*c* nous obtenons une nouvelle réalisation du théorème *a*:

T h é o r è m e (W e i e r s t r a s s). *Toute fonction réelle $f(\varphi)$ continue sur la circonférence Q est la limite uniforme d'une suite de polynômes trigonométriques (1) à coefficients réels.*

e. Choisissons sur la droite réelle une fonction réelle $g(t)$ continue et périodique de période 2π ; évidemment, on peut lui faire correspondre une fonction continue sur la circonférence Q en posant $f(\varphi) = g(\varphi + 2k\pi)$ pour tout k. Inversement, à toute fonction $f(\varphi)$ continue sur la circonférence Q, on peut faire correspondre par la formule $g(t + 2k\pi) = f(t)$ une fonction $g(t)$ continue sur toute la droite réelle. Donc, le théorème *d* se met sous la forme suivante:

T h é o r è m e. *Toute fonction réelle $g(t)$ continue et périodique de période 2π sur l'axe des t est la limite uniforme (sur tout l'axe) d'une suite de polynômes trigonométriques.*

f. La version complexe des théorèmes *a* et *e* est à un certain égard encore plus simple. En partant des formules d'Euler 8.63

$$\cos k\varphi = \frac{1}{2}(e^{ik\varphi} + e^{-ik\varphi}),$$

$$\sin k\varphi = \frac{1}{2i}(e^{ik\varphi} - e^{-ik\varphi}),$$

on peut remplacer les polynômes trigonométriques (1) par les polynômes

$$q(\varphi) = \sum_{k=-n}^{n} c_k e^{ik\varphi}. \tag{2}$$

Le fait que les polynômes (2) forment une algèbre résulte des règles de multiplication de la fonction exponentielle. La symétrie de cette algèbre découle de l'égalité $\overline{\sum c_k e^{ik\varphi}} = \sum \overline{c_k} e^{-ik\varphi}$. Deux points φ_1, φ_2 de la circonférence Q sont séparés par la fonction $e^{i\varphi}$. Le théorème 12.53*b* conduit au résultat suivant :

Théorème. *Toute fonction à valeurs complexes continue sur la circonférence Q (ou, ce qui revient au même, toute fonction continue de période 2π sur l'axe $-\infty < t < \infty$) est la limite uniforme d'une suite de polynômes trigonométriques complexes de la forme* (2).

12.55. Suites en forme de delta. Le théorème de Stone qui établit la possibilité d'approcher toute fonction continue par les fonctions d'une algèbre $B(Q)$ n'indique pourtant aucune règle de construction de fonctions approximantes. Nous signalons ici quelques méthodes concrètes d'approximations.

Puisque dans ce qui suit nous utilisons l'intégration, supposons que le compact Q est un intervalle fermé de la droite numérique ou la circonférence de rayon 1 (intervalle $[-\pi, \pi]$ aux extrémités identifiées).

a. Désignons par $U_\rho(y)$ l'intervalle ouvert de longueur 2ρ et de centre au point y. Supposons qu'il y ait, pour un point $y \in Q$ donné, une suite de fonctions non négatives $D_n(x; y)$ ($n = 1, 2, 3, \ldots$) possédant les propriétés :

1) $\int\limits_{U_\rho(y)} D_n(x; y)\, dx \xrightarrow[(n\to\infty)]{} 1$ quel que soit $\rho > 0$,

2) $\int\limits_{Q-U_\rho(y)} D_n(x; y)\, dx \xrightarrow[(n\to\infty)]{} 0$ quel que soit $\rho > 0$.

Une telle suite est dite *en forme de delta* (pour le point y). (L'origine de ce terme sera expliquée plus loin.)

b. Théorème. *Soit $D_n(x; y)$ une suite en forme de delta pour un point y; si $f(x)$ est une fonction continue par morceaux et continue au point y, alors*

$$\lim_{n\to\infty} \int_Q D_n(x; y) f(x)\, dx = f(y).$$

Démonstration. Soit $M = \sup |f(x)|$. Pour un $\varepsilon > 0$ donné, choisissons un $\delta > 0$ tel que $\rho(x, y) \leqslant \delta$ implique

$|f(x) - f(y)| \leqslant \varepsilon$. Ensuite, on a

$$\left| \int_Q D_n(x;\ y) f(x)\, dx - f(y) \right| =$$

$$= \left| \int_Q D_n(x;\ y) [f(x) - f(y)]\, dx + f(y) \left[\int_Q D_n(x;\ y)\, dx - 1 \right] \right| \leqslant$$

$$\leqslant \int_{U_\delta(y)} D_n(x;\ y) |f(x) - f(y)|\, dx + \int_{Q - U_\delta(y)} D_n(x;\ y) |f(x) -$$

$$- f(y)|\, dx + |f(y)| \left| \int_Q D_n(x;\ y)\, dx - 1 \right| \leqslant \varepsilon \int_{U_\delta(y)} D_n(x;\ y)\, dx +$$

$$+ 2M \int_{Q - U_\delta(y)} D_n(x;\ y)\, dx + M \left| \int_Q D_n(x;\ y)\, dx - 1 \right|. \qquad (1)$$

En raison des propriétés 1) et 2) d'une suite en forme de delta, la quantité obtenue est inférieure à 2ε pour n suffisamment grand, ce qu'il fallait démontrer.

c. Notons maintenant que si la fonction $D_n(x;\ y)$ est continue par rapport à l'ensemble des variables $x \in Q$, $y \in Q$, et si $f(x)$ est toujours continue par morceaux, alors

$$f_n(x) = \int_Q D_n(x;\ y) f(y)\, dy \qquad (2)$$

est une fonction continue sur Q. En effet,

$$|f_n(x') - f_n(x'')| = \left| \int_Q [D_n(x';\ y) - D_n(x'',\ y)] f(y)\, dy \right| \leqslant$$

$$\leqslant M \int_Q |D_n(x';\ y) - D_n(x'';\ y)|\, dy, \qquad (3)$$

et lorsqu'on trouve, pour un $\varepsilon > 0$ donné, un $\delta > 0$ tel que l'inégalité $|x' - x''| < \delta$ implique

$$|D_n(x';\ y) - D_n(x'';\ y)| \leqslant \frac{a}{2nM}$$

pour tout $y \in Q$, alors on a d'après (2) :

$$|f_n(x') - f_n(x'')| \leqslant \varepsilon,$$

ce qu'il nous fallait.

d. Complétons le théorème par la remarque suivante sur la convergence uniforme. Il est clair, avant tout, que si les propriétés

1) et 2) sont vérifiées pour tout point y d'un sous-ensemble $E \subset Q$ et si la fonction $f(x)$ est continue en tout point $y \in E$, alors le résultat b est valable pour tout point $y \in E$.

Nous dirons que les relations 1) et 2) ont lieu *uniformément sur un ensemble* $E \subset Q$ si, pour tout $\varepsilon > 0$, il existe un N tel que les différences du premier et du second membre des relations 1) et 2) respectivement ne dépassent pas en module le nombre ε quels que soient $n \geqslant N$ et $y \in E$.

Nous dirons qu'une fonction $f(x)$ *est uniformément continue sur* E *par rapport à* Q si, pour tout $\varepsilon > 0$, on peut trouver un $\delta > 0$ tel que $|x - y| \leqslant \delta$, $x \in Q$, $y \in E$, implique $|f(x) - f(y)| \leqslant \varepsilon$.

Alors, le théorème suivant résulte immédiatement des majorations (3):

Théorème. *Si les relations* 1) et 2), *pour tout* $\rho > 0$, *sont vérifiées uniformément sur un ensemble* E *et que la fonction* $f(x)$ *soit uniformément continue sur* E *par rapport à* Q, *alors les fonctions* $f_n(x)$ (2) *convergent uniformément sur* E *vers la fonction* $f(x)$ *lorsque* $n \to \infty$.

e. En appliquant le théorème d, on peut se servir du critère suivant de continuité uniforme d'une fonction $f(x)$ aur un ensemble E par rapport à Q:

Lemme. *Sur tout ensemble fermé* $E \subset Q$ *de points de continuité d'une fonction* $f(x)$, *cette fonction* $f(x)$ *est uniformément continue par rapport à* Q.

Démonstration. D'après le théorème de Heine (5.17b), la fonction $f(x)$ est uniformément continue sur E et, pour tout $\varepsilon > 0$, on peut trouver un $\delta_0 > 0$ tel que $|y - z| < \delta_0$, $y \in E$, $z \in E$, implique $|f(y) - f(z)| < \varepsilon/2$. Choisissons ensuite, pour tout point $y \in E$, un intervalle $|x - y| < \delta(y) \leqslant \delta_0/2$ dans lequel l'inégalité $|f(x) - f(y)| < \varepsilon/2$ soit satisfaite; puis, en appliquant 3.97, on extrait du recouvrement obtenu de l'ensemble E un recouvrement fini $|x - y_1| < \delta_1, \ldots, |x - y_n| < \delta_n$. Soit $\delta = \min(\delta_1, \ldots, \delta_n)$. Alors, quels que soient $x \in Q$, $y \in E$, $|x - y| < \delta$, on trouve un point y_k avec $|x - y_k| < \delta_k$, et l'on obtient

$$|y_k - y| < |y_k - x| + |x - y| < \delta_k + \delta \leqslant \frac{\delta_0}{2} + \frac{\delta_0}{2} = \delta_0$$

et

$$|f(x) - f(y)| \leqslant |f(x) - f(y_k)| + |f(y_k) - f(y)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ce qu'il fallait démontrer.

f. Voici une variante renforcée du théorème b: *si* $D_n(x; y)$ *est une suite en forme de delta pour un point* y *et si* $f(x)$ *est continue*

pour $x = y$, *alors, pour toute suite* $y_n \to y$, *on a*

$$\lim_{n\to\infty} \int_Q D_n(x;\ y_n) f(x)\, dx = f(y).$$

La démonstration se fait par le même calcul en précisant un peu l'estimation.

g. Considérons encore le cas où le paramètre discret n est remplacé par un paramètre continu t. Soit $D(t, x, y)$ une fonction non négative de trois variables, x et y parcourant le compact Q et t un intervalle $0 < t \leqslant b$; supposons que les conditions

1) $\lim\limits_{t\to 0} \int\limits_{|x-y|\leqslant\rho} D(t, x, y)\, dx = 1$;

2) $\lim\limits_{t\to 0} \int\limits_{|x-y|\geqslant\rho} D(t, x, y)\, dx = 0$

soient satisfaites pour tout $\rho > 0$.

Alors, *si l'on se donne, pour tout* t, *une quantité* $y(t)$ *qui tend vers a limite* y *lorsque* $t \to 0$, *on aura*

$$\lim_{t\to 0} \int_Q D(t, x, y(t)) f(x)\, dx = f(y).$$

Cette égalité se démontre par le même calcul que le théorème *b* (compte tenu de la remarque *f*). On peut dire que la fonction

$$F(t, y) = \int_Q D(t, x, y) f(x)\, dx \qquad (0 < t \leqslant b,\ y \in Q)$$

dont la définition pour $t = 0$ est donnée par la condition

$$F(0, y) = f(y)$$

est continue dans le domaine fermé $0 \leqslant t \leqslant b$, $y \in Q$.

h. On peut rejeter la condition que $D_n(x; y)$ (ou $D(t, x, y)$ dans *g*) est non négative en la remplaçant par la condition

$$\int_Q |D_n(x;\ y)|\, dx \leqslant c \quad \left(\text{ou } \int_Q |D(t, x, y)|\, dx \leqslant c\right), \tag{4}$$

où c ne dépend pas de n. La condition (4) est déjà essentielle ; sans elle le théorème tombe en défaut, ce qu'on verra dans le chapitre 14.

i. Remarque. Le terme « suite en forme de dolta » a pour origine la « fonction delta » de Dirac. P. Dirac dans son livre *The principles of quantum mechanics* (1930) définit la « fonction delta » $\delta(x)$ comme fonction sur l'axe $-\infty < x < \infty$ qui est nulle partout sauf au point $x = 0$ et qui possède la

propriété

$$\int_{-\infty}^{\infty} \delta(x)\,dx = 1. \tag{5}$$

Ensuite, il « démontre » le théorème: quelle que soit une fonction $f(x)$ continue pour $x = \xi$, on a l'égalité

$$\int_{-\infty}^{\infty} \delta(x-\xi)\,f(\xi)\,d\xi = f(x). \tag{6}$$

(La « démonstration » est bien simple: la fonction $\delta(x-\xi)$ est nulle pour $\xi \neq x$, les valeurs de $f(\xi)$ pour $\xi \neq x$ n'ont donc pas d'importance; en remplaçant $f(\xi)$ par la constante $f(x)$ et en appliquant (5) on aboutit à (6).) Il n'existe dans l'analyse classique aucune fonction possédant les propriétés imposées par Dirac, et le contenu réel de son théorème correspond à peu près au théorème *b*. Ce n'est que dans les travaux de S. Sobolev (1935) et L. Schwartz (1947) que la fonction delta est formalisée en tant qu'objet mathématique, pas comme fonction usuelle mais comme fonction généralisée (distribution) (cf., par exemple, [13]). La fonction delta de Dirac représente un exemple caractéristique de l'intuition mathématique infaillible d'un physicien qui dépasse le niveau mathématique de son temps.

12.56. Utilisation de suites en forme de delta pour la construction de fonctions approximantes.

a. Nous tenons à approcher une fonction $f(y)$ donnée par une fonction $f_n(y)$ d'une algèbre $B(Q)$. Le problème sera résolu si nous sommes en mesure de trouver une suite en forma de delta $D_n(x;\ y)$ telle que

$$f_n(y) = \int_Q D_n(x;\ y)\,f(x)\,dx \in B(Q).$$

b. Soit $Q = [0,\ 1]$ et soit $B(Q)$ l'algèbre de tous les polynômes définis sur $[0,\ 1]$. Posons, pour $n = 1,\ 2,\ \ldots$:

$$D_n(x;\ y) = C_n\,[1 - (x-y)^2]^n,$$

où

$$C_n = \frac{1}{\int_{-1}^{1} (1-t^2)^n\,dt}, \tag{1}$$

et montrons quo $D_n(x;\ y)$, pour tout $y \in (0,\ 1)$, est une suite en forme de delta. Etant donné que la fonction

$$f_n(y) = C_n \int_0^1 [1-(x-y)^2]^n\,f(x)\,dx \tag{2}$$

est un polynôme en y de degré $\leqslant 2n$ (ce qui est évident), nous obtenons l'expression pour les polynômes concrets approchant la fonction $f(y)$.

c. Lemme. *Pour tout* $\rho \in (0, 1)$, *on a*

$$\lim_{n\to\infty} \frac{\int_\rho^1 (1-t^2)^n\,dt}{\int_0^1 (1-t^2)^n\,dt} = 0.$$

La démonstration résulte des estimations simples

$$\int_\rho^1 (1-t^2)^n\,dt < (1-\rho^2)^n(1-\rho) < (1-\rho^2)^n,$$

$$\int_0^1 (1-t^2)^n\,dt > \int_0^1 (1-t)^n\,dt = \frac{1}{n+1}$$

et de la relation limite (5.58 (4) et 4.37*a*)

$$\lim_{n\to\infty} (n+1)(1-\rho^2)^n = 0.$$

En guise de conséquence, nous avons: pour tout $\rho \in (0, 1)$ on a

$$\lim_{n\to\infty} \frac{\int_0^\rho (1-t^2)^n\,dt}{\int_0^1 (1-t^2)^n\,dt} = 1.$$

d. Vérifions à présent les propriétés d'une suite en forme de delta (12.55*a* et *d*) pour la fonction $D_n(x; y)$.

En vertu du lemme nous avons pour tout $\rho \in (0, 1)$:

$$\int_{\substack{|x-y|\geqslant\rho \\ 0\leqslant x\leqslant 1}} D_n(x; y)\,dx = C_n \int_{\substack{|x-y|\geqslant\rho \\ 0\leqslant x\leqslant 1}} [1-(x-y)^2]^n\,dx =$$

$$= C_n \int_{\substack{|t|\geqslant\rho \\ -y\leqslant t\leqslant 1-y}} (1-t^2)^n\,dt \leqslant 2C_n \int_\rho^1 (1-t^2)^n\,dt =$$

$$= \frac{\int_\rho^1 (1-t^2)^n\,dt}{\int_0^1 (1-t^2)^n\,dt} \to 0 \qquad (n\to\infty),$$

ce qui démontre que la propriété 2) est satisfaite uniformément sur l'ensemble $0 \leqslant y \leqslant 1$. Ensuite, pour $y \in [\rho_0, 1-\rho_0]$, $0 < \rho < \rho_0$,

on a

$$\int\limits_{\substack{|x-y|\leqslant\rho\\0\leqslant x\leqslant 1}} D_n(x;\,y)\,dx = C_n \int\limits_{\substack{|x-y|\leqslant\rho\\0\leqslant x\leqslant 1}} [1-(x-y)^2]^n\,dx =$$

$$= C_n \int\limits_{-\rho}^{\rho} (1-t^2)^n\,dt = \frac{\int\limits_0^{\rho}(1-t^2)^n\,dt}{\int\limits_0^{1}(1-t^2)^n\,dt} \to 1 \qquad (n\to\infty),$$

ce qui démontre que la propriété 1) est satisfaite uniformément sur l'ensemble $\rho_0 \leqslant y \leqslant 1-\rho_0$.

En vertu du théorème 12.55*b* et de 12.55 *d*, les polynômes (2) forment une suite qui converge, pour tout $y \in (0, 1)$ et uniformément sur chaque intervalle $[\rho_0, 1-\rho_0]$, $\rho_0 > 0$, vers une fonction $f(y)$ continue sur $(0, 1)$.

Disons en passant que ceci démontre directement le théorème de Weierstrass pour l'intervalle $[\rho_0, 1-\rho_0]$. Par une dilatation de cet intervalle, on étend la démonstration à tout intervalle $[a, b]$.

12.57.a. On peut réaliser une construction analogue qui conduit aux polynômes approximants trigonométriques. Soient φ l'angle polaire déterminant la position d'un point sur la circonférence $Q=\{x^2+y^2=1\}$ et $B(Q)$ l'algèbre de tous les polynômes trigonométriques réels. Posons

$$D_n(\varphi;\,\psi) = C_n \cos^{2n}\frac{\varphi-\psi}{2},$$

$$C_n = \frac{1}{2\pi\int\limits_0 \cos^{2n} t\,dt} \qquad (n=1,\,2,\,\ldots) \tag{1}$$

et montrons que $D_n(\varphi;\,\psi)$ est une suite en forme de delta pour tout ψ. Comme la fonction

$$f_n(\psi) = C_n \int\limits_0^{2\pi} \cos^{2n}\frac{\varphi-\psi}{2} f(\varphi)\,d\varphi \tag{2}$$

est un polynôme trigonométrique en ψ (de degré $\leqslant 2n$), nous obtenons les polynômes trigonométriques concrets approchant $f(\varphi)$.

b. L e m m e. *Pour tout* $\rho \in (0, \pi/2)$ *on a*:

$$\lim_{n\to\infty} \frac{\int\limits_{\rho}^{\pi/2} \cos^{2n} t\,dt}{\int\limits_0^{\pi/2} \cos^{2n} t\,dt} = 0.$$

Démonstration. La fonction $\cos t$ étant décroissante pour $0 \leqslant t \leqslant \pi/2$, on a

$$\int_{\rho}^{\pi/2} \cos^{2n} t\, dt \leqslant \left(\frac{\pi}{2} - \rho\right) \cos^{2n} \rho \leqslant \frac{\pi}{2} \cos^{2n} \rho ;$$

comme la même fonction est convexe vers le haut sur l'intervalle en question, on a $\cos t \geqslant 1 - 2t/\pi$ et

$$\int_{0}^{\pi/2} \cos^{2n} t\, dt \geqslant \int_{0}^{\pi/2} \left(1 - \frac{2t}{\pi}\right)^{2n} dt = \frac{\pi}{2(2n+1)}.$$

Par conséquent, vu 5.56, on peut écrire

$$\frac{\int_{\rho}^{\pi/2} \cos^{2n} t\, dt}{\int_{0}^{\pi/2} \cos^{2n} t\, dt} \leqslant \frac{2(2n+1)}{\pi} \cdot \frac{\pi}{2} \cos^{2n} \rho \to 0$$

pour $n \to \infty$.

On obtient donc :

$$\lim_{n\to\infty} \frac{\int_{0}^{\rho} \cos^{2n} t\, dt}{\int_{0}^{\pi/2} \cos^{2n} t\, dt} = 1$$

pour tout $\rho \in (0, \pi/2)$.

c. En raison du lemme b, pour tout $\rho \in (0, \rho_0)$, $\rho_0 > 0$, on a

$$\int_{|\varphi-\psi|\geqslant\rho} D_n(\varphi ; \psi)\, d\varphi = C_n \int_{|\varphi-\psi|\geqslant\rho} \cos^{2n} \frac{\varphi-\psi}{2}\, d\varphi =$$

$$= 2C_n \int_{\frac{\pi}{2}\geqslant|t|\geqslant\frac{\rho}{2}} \cos^{2n} t\, dt = \frac{\int_{\rho/2}^{\pi/2} \cos^{2n} t\, dt}{\int_{0}^{\pi/2} \cos^{2n} t\, dt} \to 0 \qquad (n \to \infty),$$

ce qui démontre la propriété 2) d'une suite en forme de delta (12.55*a*). Ensuite, pour tout $\rho \in (0, \rho_0)$, $\rho_0 > 0$, on a

$$\int_{|\varphi-\psi|\leqslant\rho} D_n(\varphi ; \psi)\, d\varphi = C_n \int_{|\varphi-\psi|\leqslant\rho} \cos^{2n} \frac{\varphi-\psi}{2}\, d\varphi =$$

$$= 2C_n \int_{-\rho/2}^{\rho/2} \cos^{2n} t\, dt = \frac{\int_{0}^{\rho/2} \cos^{2n} t\, dt}{\int_{0}^{\pi/2} \cos^{2n} t\, dt} \to 1 \qquad (n \to \infty),$$

ce qui démontre la propriété 1). En vertu du théorème 12.55*d*, *les polynômes trigonométriques* (2) *forment une suite convergeant vers la fonction* $f(\varphi)$ *uniformément sur tout ensemble* $E \subset Q$ *sur lequel elle est uniformément continue par rapport à* Q, *en particulier* (12.55*e*), *sur tout ensemble fermé sur lequel elle est continue.*

d. Remarque. Dans les deux cas considérés, on peut estimer le degré du polynôme (algébrique ou trigonométrique) qui réalise l'approximation d'une fonction $f(x)$ à un ε donné près d'après la formule (2) ou 12.56 (2). Bien que les polynômes (2) ou 12.56 (2) aient une structure très simple, ils ne sont en général pas les meilleurs de tous les polynômes d'un degré donné. On démontre qu'il y en a, parmi les polynômes de degré n, un qui diffère d'une fonction $f(x)$ donnée et continue sur un intervalle $[a, b]$ par $12\omega\left(\frac{b-a}{n}\right)$ au plus. Ici

$$\omega(\delta) = \max_{|x-y| \leqslant \delta} |f(x) - f(y)|$$

est l'oscillation de la fonction $f(x)$ sur l'intervalle $[a, b]$ (5.17*c*). Pour les polynômes trigonométriques (sur la circonférence Q) l'estimation précédente est remplacée par $12\omega(1/n)$ (théorèmes de D. Jackson ; cf. [9]).

§ 12.6. Dérivation et intégration de fonctions à valeurs dans un espace normé

12.61. Dérivée.

a. Soit une fonction $x(t)$ définie sur un intervalle $a \leqslant t \leqslant b$, à valeurs dans un espace vectoriel normé X, réel ou complexe. Nous dirons que la fonction $x(t)$ est *dérivable en un point* $t_0 \in [a, b]$ s'il existe dans l'espace X la limite

$$x'(t_0) = \lim_{t \to t_0} \frac{x(t) - x(t_0)}{t - t_0}, \qquad (1)$$

appelée *dérivée de la fonction* $x(t)$ *au point* $t = t_0$.

b. La fonction $x(t)$ est dite *dérivable sur tout l'intervalle* $[a, b]$ si sa dérivée existe en tout point de cet intervalle ; la dérivée $x'(t)$ est alors une fonction définie sur l'intervalle $[a, b]$, à valeurs dans X.

c. Il résulte de la définition (1) que si une fonction $x(t)$ est dérivable en un point t_0, alors

$$x(t) - x(t_0) = x'(t_0)(t - t_0) + \varepsilon(t, t_0)(t - t_0),$$

où $\varepsilon(t, t_0)$ tend vers zéro dans l'espace X lorsque $t \to t_0$.

d. En particulier, la dérivabilité de $x(t)$ au point t_0 implique sa continuité en ce point. Une fonction $x(t)$ dérivable sur un intervalle $[a, b]$ y est continue.

On prouve aisément (comme dans le cas numérique) les règles principales de dérivation :

e. Si $x(t) = x_0$ est un élément constant de l'espace X, alors $x'(t) \equiv 0$.

f. Si $x(t)$ et $y(t)$ sont des fonctions dérivables à valeurs dans X, alors il en est de même de $x(t) + y(t)$, et l'on a

$$[x(t) + y(t)]' = x'(t) + y'(t).$$

g. Si $x(t)$ est une fonction dérivable à valeurs dans X et $\gamma(t)$ une fonction numérique dérivable, alors le produit $\gamma(t)\,x(t)$ est une fonction dérivable à valeurs dans X, et l'on a

$$[\gamma(t)\,x(t)]' = \gamma'(t)\,x(t) + \gamma(t)\,x'(t). \tag{2}$$

En particulier,

$$[\alpha x(t)]' = \alpha x'(t)$$

pour toute constante α.

h. Si $x(t)$ $(a \leqslant t \leqslant b)$ est une fonction dérivable de t à valeurs dans l'espace X et $t = t(\tau)$ une fonction numérique dérivable à valeurs dans l'intervalle $[a, b]$, alors $y(\tau) = x(t(\tau))$ est une fonction dérivable de τ, et l'on a

$$y'(\tau) = x'(t)\ '(\tau).$$

i. Introduisons la notion de différentielle d'une fonction $x(t)$ à valeurs dans un espace normé. Le vecteur $dx = x'(c)\,dt$, où $dt = \Delta t$ est un accroissement arbitraire du paramètre t, s'appelle *différentielle de la fonction vectorielle $x(t)$ pour $t = c$*. Ainsi, la différentielle d'une fonction est la partie linéaire principale de son accroissement dû à un accroissement de la variable t.

Le théorème sur l'invariance de la différentielle d'une fonction composée reste valable : *la différentielle de la fonction $x(t)$ a la même forme que t soit une variable indépendante ou une fonction d'une autre variable indépendante* τ (dans le dernier cas dt est la partie linéaire principale de l'accroissement de la fonction $t(\tau)$). En effet, si $g(\tau) = x[t(\tau)]$ et si $d_\tau x$ est la différentielle de la fonction x par rapport à la variable τ, alors, d'après g, on a

$$d_\tau x = g'(\tau)\,d\tau = x'(c)\,t'(\tau)\,d\tau = x'(c)\,dt = dx,$$

ce qu'il fallait démontrer.

j. Dans ce qui suit (k), nous établirons la réciproque de e : une fonction dont la dérivée est identiquement nulle est constante. Pour la raison de généralité, nous démontrerons ce théorème dans l'hypothèse que la fonction $x(t)$ n'est que dérivable par morceaux. Introduisons les définitions rigoureuses. Une fonction $x(t)$ à valeurs dans l'espace X est dite *continue par morceaux sur un intervalle fermé* $a \leqslant t \leqslant b$ s'il existe une partition $a = t_0 < t_1 < \ldots < t_n = b$ telle que $x(t)$ soit continue dans chaque intervalle (t_k, t_{k+1}) et possède les limites $x(t_k + 0)$ et $x_t(t_{k+1} - 0)$ $(k = 0, 1, \ldots, n - 1)$; comme d'habitude, aux points t_k mêmes la fonction $x(t)$ peut être définie n'importe comment ou ne pas être définie

du tout. La fonction $x(t)$ est dite *lisse par morceaux* sur $[a, b]$ si elle est continue sur $[a, b]$, possède une dérivée $x'(t)$ partout sur $[a, b]$, sauf en un nombre fini de points, et que cette dérivée soit continue par morceaux.

k. Théorème (réciproque de la propriété *e*). *Si* $x(t)$, $t \in [a, b]$, *est une fonction lisse par morceaux à valeurs dans un espace normé* X *et si la dérivée* $x'(t)$ *est nulle partout où elle existe, alors* $x(t) \equiv x_0$ (*un élément constant de l'espace* X).

Démonstration. Supposons d'abord que $x'(t) = 0$ *partout à l'intérieur de l'intervalle* $[a, b]$. Fixons un point $c \in (a, b)$ et un nombre $\varepsilon > 0$. Comme $x'(c) = 0$, il existe un voisinage du point c dans lequel on a l'inégalité

$$|x(t) - x(c)| \leqslant \varepsilon |t - c|. \tag{3}$$

Désignons par $T_\varepsilon(c)$ l'ensemble formé de tous les $t > b$ et des $t \in [c, b]$ pour lesquels l'inégalité (3) n'a pas lieu. Soit $t_0 = \inf T_\varepsilon(c)$ et supposons que $t_0 < b$. Comme $x(t)$ est continue, l'inégalité (3) qui est valable au voisinage du point t_0 le reste au point t_0 même. Comme $x'(t_0) = 0$, il existe un voisinage du point t_0 dans lequel on a l'inégalité

$$|x(t) - x(t_0)| < \frac{\varepsilon}{2} |t - t_0|. \tag{4}$$

Choisissons un $t > t_0$ pour lequel l'inégalité (4) est valable. Il résulte de (3) et (4) que

$$|x(t) - x(c)| \leqslant |x(t) - x(t_0)| + |x(t_0) - x(c)| \leqslant$$

$$\leqslant \frac{\varepsilon}{2}(t - t_0) + \varepsilon(t_0 - c) =$$

$$= \varepsilon\left(\frac{t - t_0}{2} + t_0 - c\right) < \varepsilon(t - c),$$

de sorte que le point t n'appartient pas non plus à l'ensemble $T_\varepsilon(c)$. Or, ceci contredit l'égalité $t_0 = \inf T_\varepsilon(c)$. Par conséquent, $t_0 = b$ et l'on a

$$|x(t) - x(c)| \leqslant \varepsilon(t - c)$$

pour tout $t \in [c, b]$.

Puisque ε est arbitraire, on a

$$x(t) - x(c) = 0$$

pour tout $t \in [c, b]$, donc $x(t) \equiv x(c)$.

Nous voyons que la fonction $x(t)$ est constante sur l'intervalle (c, b). Comme le point c peut être choisi autant proche que l'on veut du point a, la fonction $x(t)$ est constante sur tout l'intervalle $[a, b]$.

Considérons à présent le cas général : il y a sur l'intervalle $[a, b]$ un nombre fini de points, disons $a = c_0 < c_1 < \ldots < c_n = b$, en lesquels la fonction $x(t)$ n'a pas de dérivée ; dans chaque intervalle (c_j, c_{j+1}) $(j = 0, \ldots, n-1)$ la quantité $x'(t)$ existe et est nulle. Le raisonnement ci-dessus montre que la fonction $x(t)$ est constante sur chaque intervalle (c_j, c_{j+1}) $(j = 0, \ldots, n-1)$. La fonction $x(t)$ étant continue sur l'intervalle $[a, b]$, ses valeurs dans les intervalles voisins (c_j, c_{j+1}) et (c_{j-1}, c_j) se confondent ; il en résulte que $x(t)$ est constante sur tout l'intervalle $[a, b]$. Le théorème est démontré.

12.62. Intégration.

a. Soit $x(t)$ une fonction donnée sur un intervalle fermé $[a, b]$, à valeurs dans un espace de Banach (i.e. normé et complet) X (réel ou complexe). Etant donnée une partition

$$\Pi = \{a = t_0 \leqslant \xi_0 \leqslant t_1 \leqslant \xi_1 \leqslant t_2 \leqslant \ldots \leqslant t_{n-1} \leqslant \xi_{n-1} \leqslant t_n = b\}$$

de l'intervalle $[a, b]$, aux points marqués $\xi_0, \ldots, \xi_{n-1}$ et de paramètre $d(\Pi) = \max \Delta t_j$, on peut former la somme intégrale de Riemann

$$s_\Pi(x) = \sum_{k=0}^{n-1} x(\xi_k)\, \Delta t_k. \tag{1}$$

Bien entendu, cette somme est un élément de l'espace X. Nous affirmons que si la fonction $x(t)$ est continue par morceaux, les sommes (1) tendent, pour un morcellement illimité de la partition Π, i.e. pour $d(\Pi) \to 0$, vers une limite dans X ; nous l'appellerons *intégrale* de la fonction $x(t)$ sur l'intervalle $[a, b]$ et noterons

$$\int_a^b x(t)\, dt.$$

b. La démonstration de l'existence de l'intégrale d'une fonction continue par morceaux à valeurs dans X répète la démonstration analogue pour une fonction numérique (9.14-9.16). Indiquons ses parties principales. La fonction

$$\omega_x(\delta) = \sup_{\substack{|t'-t''| \leqslant \delta \\ t', t'' \in [a, b]}} \| x(t') - x(t'') \|$$

s'appelle *oscillation* de la fonction $x(t)$ sur l'intervalle $[a, b]$; si $x(t)$ est continue, $\omega_x(\delta)$ tend vers zéro pour $\delta \to 0$. De même que dans 9.14*c-d*, les estimations suivantes ont lieu pour les sommes intégrales de n'importe quelle fonction $x(t)$: si une partition Π s'obtient d'une autre partition Π en y ajoutant quelques points de

division, alors

$$\| s_\Pi(x) - s_{\Pi'}(x) \| \leqslant \omega_x(\delta)(b-a) \tag{2}$$

pour $d(\Pi) \leqslant \delta$; si Π et Π' sont deux partitions quelconques avec $d(\Pi) \leqslant \delta$, $d(\Pi') \leqslant \delta$, alors

$$\| s_\Pi(x) - s_{\Pi'}(x) \| \leqslant 2\omega_x(\delta)(b-a). \tag{3}$$

Les estimations (2) et (3) étant établies, il nous reste d'appliquer (pour $x(t)$ continue) la propriété $\lim_{\delta \to 0} \omega_x(\delta) = 0$ et le fait que l'espace X est complet. Le passage à une fonction continue par morceaux se réalise de même que dans 9.16.

c. Tout comme dans 9.15*c*, on peut démontrer que toute fonction $x(t)$ intégrable sur $[a, b]$ est bornée (en norme), de sorte que

$$\| x(t) \| \leqslant c.$$

Il est facile de prouver, pour les fonctions intégrables, les propriétés principales de l'intégrale :

1) $\int_a^b \alpha x(t)\,dt = \alpha \int_a^b x(t)\,dt$ (α est un nombre) ;

2) $\int_a^b [x(t) + y(t)]\,dt = \int_a^b x(t)\,dt + \int_a^b y(t)\,dt$;

3) $\int_a^b x(t)\,dt + \int_b^c x(t)\,dt = \int_a^c x(t)\,dt \quad (a < b < c)$;

4) $\left\| \int_a^b x(t)\,dt \right\| \leqslant \max_{a \leqslant t \leqslant b} \| x(t) \| (b-a)$;

5) $\left\| \int_a^b x(t)\,dt \right\| \leqslant \int_a^b \| x(t) \|\,dt.$

Elles s'obtiennent toutes en passant à la limite dans les propriétés analogues des sommes intégrales.

d. Valeur moyenne d'une fonction. De même que pour les fonctions numériques (9.15*h*), pour une fonction $x(t)$ continue par morceaux, à valeurs dans un espace de Banach X, la quantité

$$\frac{1}{b-a} \int_a^b x(t)\,dt$$

est appelée *valeur moyenne de la fonction* $x(t)$ sur l'intervalle $[a, b]$. La valeur moyenne d'une fonction $x(t)$ réelle est comprise entre ses valeurs minimale et maximale sur $[a, b]$ et est égale à une valeur $x(t_0)$ si $x(t)$ est continue.

Pour une fonction à valeurs dans un espace de Banach, même pour une fonction à valeurs complexes, la valeur moyenne peut être distincte de toute sa valeur dans l'intervalle $[a, b]$. Ainsi,

$$\frac{1}{2\pi}\int_0^{2\pi} ie^{it}\,dt = e^{it}\Big|_0^{2\pi} = 0,$$

bien que la fonction ie^{it} ne s'annule nulle part dans l'intervalle d'intégration.

e. Soit un ensemble E dans un espace vectoriel L; appelons *enveloppe convexe* de l'ensemble E l'ensemble $V(E)$ de tous les vecteurs de la forme

$$y = \sum_{k=1}^{m} \alpha_k x_k \quad (x_k \in E,\ \alpha_k \geqslant 0,\ \sum_{k=1}^{m} \alpha_k = 1,\ m = 1, 2, \ldots). \tag{4}$$

L'ensemble $V(E)$ *est convexe* (12.34*b*): en effet, si

$$\alpha \geqslant 0,\quad \beta \geqslant 0,\quad \alpha + \beta = 1,$$
$$x_k \in E,\quad y_r \in E,$$
$$x = \sum_{k=1}^{m} \alpha_k x_k \in V(E),\quad y = \sum_{r=1}^{n} \beta_r y_r \in V(E),$$

alors le vecteur

$$\alpha x + \beta y = \alpha \sum_{k=1}^{m} \alpha_k x_k + \beta \sum_{r=1}^{n} \beta_r y_r = \sum_{k=1}^{m} \alpha\cdot\alpha_k x_k + \sum_{r=1}^{n} \beta\cdot\beta_r y_r$$

appartient lui aussi à $V(E)$, car $\alpha\alpha_k \geqslant 0$, $\beta\beta_r \geqslant 0$ et

$$\sum_{1}^{m} \alpha\cdot\alpha_k + \sum_{1}^{n} \beta\cdot\beta_r = \alpha \sum_{1}^{m} \alpha_k + \beta \sum_{1}^{n} \beta_r = \alpha + \beta = 1.$$

D'autre part, *tout ensemble convexe* P *contenant un ensemble donné* E *contient également tous les vecteurs de la forme* (4). Pour $m = 2$, cela découle de la définition même d'un ensemble convexe. Raisonnons par récurrence: supposons que cette affirmation soit juste pour n'importe quels $m - 1$ vecteurs, et montrons qu'elle est alors juste pour m vecteurs quelconques $x_1, \ldots, x_m \in E$. Nous avons

$$z = \alpha_1 x_1 + \ldots + \alpha_m x_m =$$
$$= \frac{\alpha_1 x_1 + \ldots + \alpha_{m-1} x_{m-1}}{\alpha_1 + \ldots + \alpha_{m-1}} (\alpha_1 + \ldots + \alpha_{m-1}) + \alpha_m x_m =$$
$$= (\alpha_1 + \ldots + \alpha_{m-1})\, z_1 + \alpha_m x_m.$$

Le vecteur z_1 appartient à l'ensemble P d'après la supposition de récurrence; le vecteur z y appartient comme point du segment liant z_1 et x_m.

On peut donc dire que l'ensemble $V(E)$ que nous avons construit est le plus petit ensemble convexe contenant E. Si E est lui-même convexe, alors, évidemment, $V(E) = E$.

f. Dans un espace de Banach X, ce n'est pas tout ensemble convexe qui est fermé (un intervalle ouvert de la droite en est un exemple). Ayant donné un ensemble $E \subset X$, on peut former son enveloppe convexe $V(E)$, puis sa fermeture $\overline{V(E)}$; cette dernière s'appelle *enveloppe convexe fermée de l'ensemble E*. L'ensemble $\overline{V(E)}$ est encore convexe; en général, *la fermeture d'un ensemble convexe est un ensemble convexe*, puisqu'il résulte de

$$x = \lim x_n, \quad y = \lim y_n, \quad x_n \in V, \quad y_n \in V,$$

que

$$\alpha x + \beta y = \lim (\alpha x_n + \beta y_n) \in \overline{V}.$$

L'ensemble $\overline{V(E)}$ est le plus petit ensemble convexe fermé contenant l'ensemble donné E.

g. Théorème. *La moyenne* (*d*) *d'une fonction* $x(t)$ *continue par morceaux à valeurs dans un espace de Banach* X *appartient à l'enveloppe convexe fermée de l'ensemble des valeurs de* $x(t)$ *sur l'intervalle* $[a, b]$.

La démonstration résulte de la définition de la moyenne

$$\frac{1}{b-a}\int_a^b x(t)\,dt = \frac{1}{b-a}\lim_{d(\Pi)\to 0}\sum_{k=1}^{n} x(\xi_k)\,\Delta t_k,$$

car la somme intégrale à droite appartient à l'enveloppe convexe des valeurs de la fonction (parce que $\frac{1}{b-a}\sum_{k=1}^{n}\Delta t_k = 1$).

Pour l'exemple donné dans *d*, la moyenne de la fonction te^{it} sur $[0, 2\pi]$ qui vaut 0 appartient à l'enveloppe convexe de toutes les valeurs de la fonction ie^{it} sur $[0, 2\pi]$: ces valeurs remplissent la circonférence de rayon 1, leur enveloppe convexe est tout le cercle limité par cette circonférence.

h. Intégrales impropres. La théorie des intégrales impropres des fonctions à valeurs dans un espace de Banach peut être construite par analogie avec le cas des fonctions numériques (chapitre 11). Indiquons-en les étapes principales. Soit $x(t)$ une fonction à valeurs dans un espace de Banach X, définie sur la demi-droite $a \leqslant t < \infty$ et intégrable (par exemple, continue par morceaux) sur tout intervalle $a \leqslant t \leqslant b$. L'intégrale impropre de

premier espèce

$$\int_a^\infty x(t)\,dt \tag{5}$$

est définie comme limite (pour la norme de l'espace X) de l'intégrale

$$\int_a^b x(t)\,dt \tag{6}$$

pour $b \to \infty$, à condition que la limite existe.

En particulier, si l'intégrale impropre usuelle

$$\int_a^\infty \|x(t)\|\,dt \tag{7}$$

existe, il en est de même de l'intégrale impropre (5) qui est alors dite *absolument convergente*; de plus, on a l'estimation suivante:

$$\left\| \int_a^\infty x(t)\,dt \right\| \leqslant \int_a^\infty \|x(t)\|\,dt. \tag{8}$$

L'existence de l'intégrale (5) dans l'hypothèse d'existence de l'intégrale (7) résulte du critère de Cauchy: *pour que l'intégrale* (5) *existe, il faut et il suffit que, pour tout* $\varepsilon > 0$, *il existe un* N *tel que l'inégalité*

$$\left\| \int_p^q x(t)\,dt \right\| < \varepsilon$$

a lieu quels que soient $p \geqslant N$, $q \geqslant N$.

Les définitions des intégrales impropres de deuxième et de troisième espèce sont généralisées de façon analogue.

12.63. Intégrale et primitive.

a. Soit $x(t)$ une fonction continue par morceaux sur un intervalle $]a, b]$ à valeurs dans un espace de Banach X; montrons que la *fonction*

$$F(t) = \int_a^t x(\xi)\,d\xi \tag{1}$$

a pour dérivée, en tout point $t = t_0$ *de continuité de la fonction* $x(t)$, *la valeur* $x(t_0)$.

Selon les règles d'intégration 12.62*c*, nous avons

$$\frac{F(t)-F(t_0)}{t-t_0}=\frac{1}{t-t_0}\int_{t_0}^{t}x(\xi)\,d\xi=$$

$$=\frac{1}{t-t_0}\int_{t_0}^{t}x(t_0)\,d\xi+\frac{1}{t-t_0}\int_{t_0}^{t}[x(\xi)-x(t_0)]\,d\xi=$$

$$=x(t_0)+\frac{1}{t-t_0}\int_{t_0}^{t}[x(\xi)-x(t_0)]\,d\xi.$$

Ensuite, en raison de la continuité de $x(t)$ au point t_0, on a

$$\left|\frac{1}{t-t_0}\int_{t_0}^{t}[x(\xi)-x(t_0)]\,d\xi\right|\leqslant\max_{\xi\in[t_0,\,t]}\|x(\xi)-x(t_0)\|\to 0$$

pour $t\to t_0$, d'où le résultat cherché.

b. Une fonction $G(t)$ à valeurs dans l'espace de Banach X est appelée *primitive* de la fonction $x(t)$ continue par morceaux si $G'(t)=x(t)$ en tout point de continuité de $x(t)$. S'il y a deux primitives $G(t)$ et $F(t)$ de la fonction $x(t)$, alors

$$[G(t)-F(t)]'=G'(t)-F'(t)=x(t)-x(t)=0,$$

et, en vertu du théorème 12.61*k*, la fonction $G(t)-F(t)$ est constante. Nous voyons que la *différence de deux primitives est un élément constant de l'espace* X. Puisque la fonction (1) est, on l'a vu, l'une des primitives, toute primitive a la forme

$$G(t)=\int_{a}^{t}x(\xi)\,d\xi+x_0,$$

x_0 étant un élément fixe de l'espace X. En particulier, pour toute primitive on a la formule

$$G(b)-G(a)=\int_{a}^{b}x(\xi)\,d\xi,$$

généralisant la formule de Newton-Leibniz.

c. Inversement, *soit $G(t)$ une fonction dérivable de la variable $t\in[a,\,b]$ ayant la dérivée continue par morceaux; alors, pour tout t, on a l'égalité*

$$G(t)=G(a)+\int_{a}^{t}G'(\xi)\,d\xi. \tag{2}$$

En effet, désignons provisoirement par $G^*(t)$ le second membre de (2). Cette fonction, d'après *a*, a pour dérivée la fonction $G'(t)$

en tout point de continuité de cette dernière. La fonction $G(t)$ possède la même propriété, donc, les deux fonctions étant continues, nous avons $G^*(t) - G(t) = c_0 = \text{const}$ d'après b. Or, $G^*(a) = G(a)$, d'où $c_0 = 0$ et la formule (2) est établie.

d. Pour les fonctions numériques dérivables, nous avons la formule de Lagrange (7.44)

$$G(b) - G(a) = (b - a)\,Q,$$

où Q est un nombre compris entre la plus grande et la plus petite valeur de la fonction $G'(t)$ sur $[a, b]$, autrement dit la valeur de la fonction $G'(t)$ en un point $t = t_0$. Pour une fonction $G(t)$ dérivable à valeurs dans l'espace de Banach X, cette formule reste valable, à ceci près que le point Q appartient cette fois à l'enveloppe convexe fermée de l'ensemble des valeurs de $G'(t)$ sur $[a, b]$. Ce fait résulte directement de 12.62g et de la formule (2).

e. La formule de Newton-Leibniz a pour conséquence, tout comme dans 9.51a, la formule d'intégration par parties:

$$\int_a^b u(t)\,dv(t) = u(t)\,v(t)\Big|_a^b - \int_a^b v(t)\,du(t).$$

Ici l'une des fonctions $u(t)$, $v(t)$ est numérique, l'autre vectorielle (à valeurs dans l'espace X), les deux étant lisses par morceaux.

f. De même que dans 9.54, on obtient la formule d'intégration par substitution

$$\int_{\tau=\alpha}^{\beta} x(t(\tau))\,t'(\tau)\,d\tau = \int_{t=a}^{b} x(t)\,dt$$

dans les mêmes hypothèses sur les fonctions $x(t)$ et $t(\tau)$ et les nombres α, β, a, b.

12.64. Dérivées d'ordre supérieur, différentielles d'ordre supérieur, formule de Taylor.

a. Les dérivées supérieures d'une fonction $x(t)$ à valeurs dans l'espace X sont définies, comme dans le cas d'une fonction numérique, par récurrence. *La dérivée* n-ième est, par définition, la dérivée première de la dérivée d'ordre $n-1$ si cette dernière est une fonction dérivable pour $a \leqslant t \leqslant b$. Toutes les dérivées ainsi obtenues sont toujours des fonctions vectorielles à valeurs dans le même espace X.

Les dérivées d'ordre supérieur d'une fonction vectorielle sont désignées de même que celles d'une fonction numérique:

$$(x'(t))' \equiv x''(t),\ (x''(t))' \equiv x'''(t),\ \ldots,\ (x^{(n)}(t))' \equiv x^{(n+1)}(t).$$

b. Les différentielles d'ordre supérieur sont définies elles aussi par récurrence :

$$d^2x(t) \equiv d[dx(t)] \equiv d[x'(t)\,dt] = x''(t)\,dt^2,$$

. . . . . .

$$d^{n+1}x(t) \equiv d[d^n x(t)] \equiv d[x^{(n)}(t)\,dt^n] = x^{(n+1)}(t)\,dt^{n+1}.$$

Contrairement à la différentielle première, les différentielles d'ordre supérieur changent de forme lorsqu'on passe à une nouvelle variable indépendante (à l'exception du changement linéaire de variable).

c. Si toutes les dérivées de la fonction $x(t)$, y compris la $(n+1)$-ième, existent pour $a \leqslant t \leqslant b$, alors on a la formule de Taylor

$$\Delta x(a) \equiv x(b) - x(a) =$$

$$= \left\{ \begin{array}{l} dx(a) + \frac{1}{2} d^2x(a) + \ldots + \frac{1}{n!} d^n x(a), \\ x'(a)(b-a) + \frac{(b-a)^2}{2} x''(a) + \ldots + \frac{(b-a)^n}{n!} x^{(n)}(a) \end{array} \right\} + Q_n,$$

avec le reste que l'on peut mettre sous la forme

$$Q_n = \frac{1}{n!} \int_a^b x^{(n+1)}(t)\,(b-t)^n\,dt.$$

La démonstration de la formule de Taylor se fait par la même méthode que dans 9.52*a*, en utilisant la formule d'intégration par parties 12.63*c*. En partant de l'expression pour le reste on établit les estimations

$$\| Q_n \| \leqslant \max_{a \leqslant t \leqslant b} \| x^{(n+1)}(t) \| \frac{1}{n!} \int_a^b (b-t)^n\,dt =$$

$$= \max_{a \leqslant t \leqslant b} \| x^{(n+1)}(t) \| \frac{(b-a)^{n+1}}{(n+1)!}.$$

12.65. Suites et séries de fonctions à valeurs dans X.

a. Soit $x_1(t)$, $x_2(t)$, . . ., $x_n(t)$. . . . une suite de fonctions de la variable $t \in [a, b]$, à valeurs dans l'espace de Banach X. Par définition, une fonction $x(t)$ est la *limite* de la suite $x_n(t)$ pour $n \to \infty$ si la relation

$$\lim_{n \to \infty} \| x(t) - x_n(t) \| = 0$$

est remplie pour tout $t \in [a, b]$. La suite $x_n(t)$ est dite *uniformément convergente* vers la limite $x(t)$ si

$$\lim_{n\to\infty} \sup_t \| x(t) - x_n(t) \| = 0,$$

autrement dit, si, pour tout $\varepsilon > 0$, il existe un numéro N tel que $n \geqslant N$ implique $\| x(t) - x_n(t) \| \leqslant \varepsilon$ quel que soit $t \in [a, b]$. Nous avons déjà vu dans 5.96 que la limite d'une suite uniformément convergente de fonctions *continues* est encore une fonction continue. Ont lieu les analogues des théorèmes 9.72 et 9.77 qui étaient démontrés pour les fonctions à valeurs numériques, à savoir :

b. Théorème. *Si une suite $x_n(t)$ de fonctions intégrables converge uniformément sur $[a, b]$ vers une fonction $x(t)$, alors $x(t)$ est elle aussi intégrable et l'on a*

$$\lim_{n\to\infty} \int_a^\tau x_n(t)\,dt = \int_a^\tau x(t)\,dt$$

uniformément par rapport à $\tau \in [a, b]$. En particulier,

$$\lim_{n\to\infty} \int_a^b x_n(t)\,dt = \int_a^b x(t)\,dt.$$

c. Théorème. *Si une suite $x_n(t)$ de fonctions lisses par morceaux converge pour au moins un point $t_0 \in [a, b]$ et la suite $x_n'(t)$ de leurs dérivées converge uniformément sur $[a, b]$ vers une fonction $g(t)$ continue par morceaux, alors la suite $x_n(t)$ converge uniformément sur $[a, b]$ vers une fonction $x(t)$ lisse par morceaux et $x'(t) = \lim_{n\to\infty} x_n'(t) = g(t)$ aux points de continuité de $g(t)$.*

Les démonstrations de ces théorèmes imitent celles des théorèmes 9.72 et 9.77.

d. Une série

$$x_1(t) + x_2(t) + \ldots + x_n(t) + \ldots \qquad (1)$$

de fonctions à valeurs dans l'espace X est dite *convergente* sur un intervalle $[a, b]$ si la suite de ses sommes partielles

$$s_1(t) = x_1(t), \ \ldots, \ s_n(t) = x_1(t) + \ldots + x_n(t), \ \ldots$$

converge quel que soit $t \in [a, b]$; la limite de la suite $s_n(t)$ est appelée *somme* de la série (1). La série (1) est dite *uniformément convergente* sur $[a, b]$ si la suite $s_n(t)$ converge uniformément. Les théorèmes *b* et *c* ont pour conséquences certaines conditions suffisantes d'intégrabilité terme à terme et de dérivabilité d'une série de fonctions ; les énoncés de ces conditions sont laissés au lecteur.

12.66. Fonctions analytiques. Soit $x(\zeta)$ une fonction à valeurs dans un espace complexe normé X, définie dans un domaine G du plan complexe $\zeta = \xi + i\eta$. Cette fonction est dite *dérivable en un point* $\zeta_0 \in G$ s'il existe dans l'espace X un élément

$$x'(\zeta_0) = \lim_{h \to 0} \frac{x(\zeta_0 + h) - x(\zeta_0)}{h}$$

appelé *dérivée de la fonction* $x(\zeta)$ *par rapport à la variable complexe* ζ *au point* ζ_0. La fonction $x(\zeta)$ est dite *analytique dans le domaine* G si elle est dérivable par rapport à ζ en tout point $\zeta_0 \in G$.

Pour les fonctions analytiques à valeurs dans l'espace X, les propositions de la théorie ordinaire des fonctions analytiques (chapitre 10) restent valables. La définition de l'intégrale le long d'un ligne du plan complexe, qui est nécessaire pour l'établissement de la théorie, est formulée de la façon habituelle comme suit. Soit L un chemin lisse par morceaux dans le domaine G: $\zeta = \zeta(t)$, où t parcourt un intervalle $a \leqslant t \leqslant b$. Soient ensuite $\Pi = \{a = t_0 < t_1 < \ldots < t_n = b\}$ une partition de l'intervalle $[a, b]$, $\zeta_j = \zeta(t_j)$ $(j = 0, 1, \ldots, n)$ les points correspondants du chemin L et $\Delta\zeta_j = \zeta_{j+1} - \zeta_j$; posons

$$\int_L x(\zeta)\, d\zeta = \lim_{d(\Pi) \to 0} \sum_{j=0}^{n-1} x(\zeta_j)\, \Delta\zeta_j.$$

L'existence de cette intégrale pour toute fonction continue par morceaux à valeurs dans un espace normé complet X se démontre de même que pour une fonction numérique (10.21). On démontre également, pour une fonction analytique $x(\zeta)$, le théorème de Cauchy: *si une fonction* $x(\zeta)$ *est analytique dans un domaine simplement connexe* G, *alors, pour tout contour fermé* L *inclus dans le domaine* G, *on a*

$$\oint_L x(\zeta)\, d\zeta = 0.$$

En partant du théorème de Cauchy on établit de la façon ordinaire la formule de Cauchy

$$x(\zeta_0) = \frac{1}{2\pi i} \oint_L \frac{x(\zeta)\, d\zeta}{\zeta - \zeta_0} \quad (\zeta_0 \text{ à l'intérieur de } L),$$

puis les autres propositions du § 10.3. En particulier, une fonction analytique $x(\zeta)$ possède, dans le domaine G, les dérivées de tous les ordres et se développe, dans tout cercle $Q = \{|\zeta - \zeta_0| < \rho\}$ inclus dans le domaine G, en sa série de Taylor

$$x(\zeta) = \sum_{m=0}^{\infty} a_m (\zeta - \zeta_0)^m, \tag{1}$$

où $a_m = \frac{1}{m!} x^{(m)}(\zeta_0)$ $(m = 0, 1, 2, \ldots)$.

Le rayon de convergence de cette série est égal à la distance du point ζ_0 à la singularité la plus proche de la fonction $x(\zeta)$ (i.e. au point en lequel la fonction $x(\zeta)$ cesse d'être dérivable) et peut être trouvé d'après la formule de Cauchy-Hadamard (12.39*l*)

$$\frac{1}{R} = \overline{\lim_{m\to\infty}} \sqrt[m]{\|a_m\|}.$$

Les dérivées successives de la fonction $x(\zeta)$ s'obtiennent en dérivant terme à terme la série (1):

$$x'(\zeta) = \sum_{m=1}^{\infty} m a_m (\zeta - \zeta_0)^{m-1},$$

. . . . . .

$$x^{(k)}(\zeta) = \sum_{m=k}^{\infty} m(m-1)\ldots(m-k+1)\, a_m (\zeta - \zeta_0)^{m-k},$$

. . . . . .

§ 12.7. Opérateurs linéaires continus

12.71. Nous avons donné la définition d'un opérateur linéaire dans 12.15. Il était dit qu'une application A d'un espace vectoriel X dans un espace vectoriel Y (sur un même corps K) s'appelle *opérateur linéaire* si les conditions

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2$$

sont satisfaites pour tous x_1 et x_2 de l'espace X quels que soient les nombres α_1, α_2 du corps K. Si l'espace Y est unidimensionnel et $Y = K$, l'opérateur A s'appelle *fonctionnelle linéaire*.

Ici nous considérons les opérateurs linéaires d'un espace normé X dans un espace normé Y, les deux étant réels pour le moment.

a. Conformément à la définition générale d'une fonction continue 5.11*a*, un opérateur linéaire A d'un espace normé X dans un espace normé Y est dit *continu* pour $x = x_0 \in X$ si, quel que soit $\varepsilon > 0$, il existe un $\delta > 0$ tel que $|x - x_0| \leqslant \delta$ implique $|Ax - Ax_0| \leqslant \varepsilon$. Comme d'ordinaire, il y a une définition équivalente: l'opérateur A est continu pour $x = x_0$ si $Ax_n \to Ax_0$ (dans Y) dès que $x_n \to x_0$ (dans X).

b. Un opérateur linéaire A d'un espace X dans un espace Y est dit *borné* s'il est borné sur la boule unité de l'espace X, de sorte que $|x| \leqslant 1$ implique $|Ax| \leqslant c$ avec une constante fixe c. Dans ce cas, la quantité

$$\|A\| = \sup_{|x|\leqslant 1} |Ax|$$

s'appelle *norme de l'opérateur* A. Pour tout vecteur $x \in X$, on a $\left|\frac{x}{|x|}\right| = 1$, d'où $\left|A\frac{x}{|x|}\right| \leqslant \|A\|$ et, par conséquent,

$$|Ax| \leqslant \|A\| \, |x|. \tag{1}$$

c. *Si un opérateur linéaire* A *est borné, il est continu en tout point* x_0 *de l'espace* X.

Démonstration. Soient A un opérateur borné et $\|A\|$ sa norme. Alors on a

$$|Ax - Ax_0| = |A(x - x_0)| \leqslant \|A\| \, |x - x_0| < \varepsilon$$

pour un $\varepsilon > 0$ donné et pour $|x - x_0| < \varepsilon/\|A\|$.

d. *Si un opérateur linéaire* A *est continu pour au moins un point* $x = x_0$, *alors il est borné.*

Démonstration. Trouvons un δ de façon à avoir $|Ax - Ax_0| \leqslant 1$ dès que $|x - x_0| \leqslant \delta$. Soient $|z| \leqslant 1$ et $x = x_0 + \delta z$. Alors on a

$$|x - x_0| = \delta |z| \leqslant \delta,$$

$$|Ax - Ax_0| = |A(x - x_0)| = \delta |Az| \leqslant 1,$$

$$|Az| \leqslant \frac{1}{\delta},$$

ce qu'il fallait démontrer.

e. En tant que conséquence, nous avons: *un opérateur linéaire qui est continu pour au moins un point de l'espace* X *est continu en tout son point.*

Les trois théorèmes suivants sont également valables pour un opérateur continu A d'un espace de Banach X dans un espace de Banach Y:

f. *Si une série* $\sum_1^\infty x_n = s$ *converge dans l'espace* X, *alors*

$$\sum_1^\infty Ax_n = As.$$

g. *Si* $x(t)$ *est une fonction continue par morceaux sur un intervalle* $a \leqslant t \leqslant b$, *à valeurs dans l'espace* X, *alors on a*

$$A\left\{\int_a^b x(t)\,dt\right\} = \int_a^b [Ax(t)]\,dt.$$

h. *Si* $x(t)$ *est une fonction dérivable pour* $t = t_0$, *à valeurs dans l'espace* X, *alors on a*

$$A[x'(t_0)] = (Ax)'(t_0).$$

La démonstration des trois théorèmes ci-dessus suit une même voie. Somme d'une série, intégrale et dérivée sont les résultats de

certains opérations linéaires et passages à la limite, or un opérateur linéaire continu commute avec les opérations linéaires ainsi qu'avec un passage à la limite; il commute donc avec les résultats définitifs.

i. Si trois opérateurs A, A_1, A_2 d'un espace vectoriel normé X dans un espace vectoriel normé Y sont bornés, il en est de même des opérateurs $A_1 + A_2$ et αA (12.15*e*) pour tout α réel puisque, pour $|x| \leqslant 1$, on a

$$|(A_1 + A_2)x| = |A_1x + A_2x| \leqslant |A_1x| + |A_2x| \leqslant$$
$$\leqslant \|A_1\| + \|A_2\|,$$
$$|\alpha Ax| = |\alpha|\,|Ax| \leqslant |\alpha|\,\|A\|.$$

De plus, les formules écrites montrent que

$$\|A_1 + A_2\| = \sup_{|x|\leqslant 1} |(A_1 + A_2)x| \leqslant \|A_1\| + \|A_2\|,$$
$$\|\alpha A\| = \sup_{|x|\leqslant 1} |\alpha Ax| = |\alpha| \sup_{|x|\leqslant 1} |Ax| = |\alpha|\,\|A\|$$

On peut donc dire que l'espace L (X, Y) des opérateurs linéaires bornés de X dans Y est un *espace normé* avec la norme 12.71*b*:

$$\|A\| = \sup_{|x|\leqslant 1} |Ax|.$$

j. Soient B un opérateur linéaire borné d'un espace normé X dans un espace normé Y et A un opérateur linéaire borné de Y dans un espace normé Z. Alors l'opérateur linéaire P = AB de X dans Z est défini (12.15*g*). Montrons que l'opérateur P est lui aussi borné. En effet, pour tout $x \in X$ on a

$$|ABx| \leqslant \|A\|\,|Bx| \leqslant \|A\|\,\|B\|\,|x|,$$

d'où l'on voit que P = AB est un opérateur borné et que

$$\|AB\| \leqslant \|A\|\,\|B\|. \tag{2}$$

k. En particulier, pour un opérateur A dans X, on a

$$\|A^2\| = \|AA\| \leqslant \|A\|^2$$

et de même

$$\left.\begin{aligned} &\|A^3\| = \|A^2A\| \leqslant \|A^2\|\,\|A\| \leqslant \|A\|^3, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &\|A^n\| \leqslant \|A^{n-1}A\| \leqslant \|A^{n-1}\|\,\|A\| = \|A\|^n. \end{aligned}\right\} \tag{3}$$

l. En guise d'exemple, trouvons la norme d'un opérateur linéaire spécial défini dans l'espace $R^*[a, b]$ des fonctions réelles continues sur l'intervalle $a \leqslant t \leqslant b$. Soit $D(t, \lambda)$, pour toute valeur du paramètre λ d'un ensemble Λ, une fonction réelle continue de $t \in [a, b]$.

et supposons que la quantité

$$D = \sup_{\lambda} \int_a^b |D(t, \lambda)|\, dt$$

soit finie. Pour $x(t) \in R^1[a, b]$ posons

$$y(\lambda) \equiv A[x] = \int_a^b D(t, \lambda)\, x(t)\, dt. \tag{4}$$

L'opérateur A transforme toute fonction $x(t)$ en une fonction $y(\lambda)$ définie sur l'ensemble Λ. La fonction $y(\lambda)$ est bornée car

$$|A(x)| = |y(\lambda)| = \left| \int_a^b D(t, \lambda)\, x(t)\, dt \right| \leqslant$$

$$\leqslant \max_{a \leqslant t \leqslant b} |x(t)| \cdot \int_a^b |D(t, \lambda)|\, dt \leqslant D \|x\|. \tag{5}$$

Ainsi, la formule (4) définit un opérateur de l'espace $R^s[a, b]$ dans l'espace $R(\Lambda)$ des fonctions réelles bornées $y(\lambda)$. Ce dernier espace est muni de la norme naturelle

$$\|y\| = \sup_{\lambda \in \Lambda} |y(\lambda)|.$$

L'opérateur A est évidemment linéaire; il résulte de l'inégalité (5) qu'il est borné et que sa norme ne dépasse pas la quantité D. Montrons que $\|A\| = D$.

Considérons la fonction $x_n(t, \lambda) = u_n[D(t, \lambda)]$, où $u_n(\tau)$ est une fonction continue qui vaut -1 pour $\tau \leqslant -1/n$, $+1$ pour $\tau \geqslant +1/n$ et qui est linéaire dans l'intervalle $[-1/n, 1/n]$ (fig. 12.10). La fonction $x_n(t, \lambda)$ est elle aussi continue en t; le produit $D(t, \lambda) x_n(t, \lambda)$ est une fonction non négative valant $|D(t, \lambda)|$ pour $|D(t, \lambda)| \geqslant 1/n$ et ne dépassant pas $|D(t, \lambda)|$ aux autres points. Pour un $\lambda \in \Lambda$ fixe, la fonction $x_n(t, \lambda)$ est un élément de l'espace $R^s(a, b)$. De plus

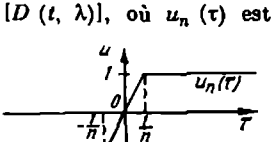


Fig. 12.10.

$$A[x_n(t, \lambda)] \geqslant \int_{|D(t, \lambda)| \geqslant 1/n} |D(t, \lambda)|\, dt \geqslant$$

$$\geqslant \int_a^b |D(t, \lambda)|\, dt - \frac{1}{n}(b-a).$$

Comme $\|x_n(t, \lambda)\| \leqslant 1$, on a

$$\|A\| = \sup_{\|x\| \leqslant 1} \|Ax(t)\| \geqslant \sup_{n, \lambda} |A[x_n(t, \lambda)]| = \sup_{\lambda} \int_a^b |D(t, \lambda)|\, dt.$$

Compte tenu de l'inégalité (5), nous obtenons

$$\|A\| = \sup_{\lambda \in \Lambda} \int_a^b |D(t, \lambda)|\, dt,$$

ce qu'il fallait démontrer.

1. Soit $D(t)$ une fonction continue de $t \in [a, b]$; alors la formule

$$F[x] = \int_a^b D(t)\, x(t)\, dt \tag{6}$$

définit une fonctionnelle linéaire dans l'espace $R^*(a, b)$ que l'on peut considérer comme cas particulier de l'opérateur décrit dans l, l'ensemble des valeurs du paramètre λ étant formé d'un seul point. En appliquant le résultat l on obtient : la norme de la fouctionnelle (6) vaut

$$\|F\| = \int_a^b |D(t)|\, dt.$$

12.72. Théorème sur l'application ouverte.

a. Soit $y = f(x)$ une fonction définie sur un ensemble X à valeurs dans un ensemble Y. Tous les points $y = f(x)$, où x parcourt un sous-ensemble $Q \subset X$, forment l'*image* du sous-ensemble Q qui se note $f(Q)$. L'ensemble de tous les points $x \in X$ pour lesquels $y = f(x)$ appartient à un sous-ensemble $F \subset Y$ s'appelle *image réciproque* du sous-ensemble F et se note $f^{-1}(F)$.

Si X et Y sont des espaces métriques et $y = f(x)$ une fonction continue, alors l'*image réciproque* $f^{-1}(G)$ *de tout sous-ensemble ouvert* $G \subset Y$ *est un sous-ensemble ouvert dans* X (5.14*a*).

Cependant, l'*image* $f(G)$ d'un ensemble ouvert $G \subset X$ n'est pas forcément un ensemble ouvert dans Y. Par exemple, si X est la droite $-\infty < x < \infty$ et Y la droite $-\infty < y < \infty$, la fonction $y = f(x)$ étant constante, alors l'image de tout ensemble ouvert (et, en général, de tout ensemble $G \subset X$) se réduit à un seul point y qui ne constitue pas un ensemble ouvert dans Y. Si l'on renforçait l'hypothèse en exigeant que la fonction $f(x)$ applique l'espace X sur Y, alors on considérerait la fonction continue valant $(x - 1)^3$ pour $x \geqslant 1$, $(x + 1)^3$ pour $x \leqslant -1$ et 0 pour $|x| < 1$; cette fonction qui applique l'axe X tout entier sur l'axe Y transforme l'ensemble ouvert $\{|x| < 1\}$ toujours en un seul point $y = 0$.

Supposons que la fonction continue $y = f(x)$ applique d'une façon *bijective* l'espace X dans l'espace Y. Choisissons pour X l'es-

pace $D_1(a, b)$ des fonctions $x(t)$ continûment dérivables sur l'intervalle $[a, b]$ (12.25) muni de sa métrique naturelle et pour Y le sous-ensemble de l'espace $R^1(a, b)$ de toutes les fonctions continues sur $[a, b]$ (toujours avec sa métrique naturelle) qui est formé des fonctions continûment dérivables; nous avons le droit de considérer ce sous-ensemble comme espace métrique. Considérons l'application $y = f(x)$ qui fait correspondre toute fonction $x = x(t) \in D_1(a, b)$ à elle-même: $y = y(t) \equiv x(t) \in R^1(a, b)$. Cette application est continue car la convergence $x_n(t) \to x(t)$ dans $D_1(a, b)$ implique, bien entendu, la convergence $y_n(t) = x_n(t) \to y(t) \equiv x(t)$ dans $R^1(a, b)$. L'application $y = f(x)$ est évidemment bijective. Néanmoins, l'image d'un ensemble ouvert dans X, par exemple de la boule unité ouverte V dans $D_1(a, b)$, n'est pas ouverte dans Y, parce que tout voisinage d'un point $y_0(t) \in f(V)$ défini par l'inéquation $\max |y(t) - y_0(t)| < \varepsilon$ contient des fonctions à dérivée indéfiniment grande.

b. Nous comprenons maintenant que les hypothèses du théorème suivant sont essentielles:

Théorème sur l'application ouverte (Banach). *Soit* A *un opérateur linéaire continu appliquant d'une façon bijective un espace normé et complet* X *sur un espace normé et complet* Y. *Alors l'opérateur* A *transforme tout ensemble ouvert* $G \subset X$ *dans un ensemble ouvert* $f(G) \subset Y$.

Démonstration. Désignons par V_r la boule $\{x: |x| < r\}$. Nous allons d'abord démontrer que la fermeture de l'ensemble $A(V_1)$ dans Y contient une boule de l'espace Y.

Par hypothèse, on a $Y = A(X) = A\left(\bigcup_{n=1}^{\infty} V_n\right) = \bigcup_{n=1}^{\infty} A(V_n)$. D'autant plus, $Y = \bigcup_{n=1}^{\infty} \overline{A(V_n)}$. En vertu du théorème de Baire (3.75*a*), il existe un numéro $n = N$ tel que l'ensemble $\overline{A(V_N)}$ contienne une boule $\{y: |y - y_0| < \varepsilon\}$. Comme l'ensemble $\overline{A(V_N)}$ est évidemment équilibré, il contient également la boule $\{y: |y + y_0| < \varepsilon\}$. De plus, l'ensemble $\overline{A(V_N)}$ est convexe (puisqu'un opérateur linéaire transforme un ensemble convexe en un ensemble convexe, et la fermeture d'un ensemble convexe est convexe d'après 12.62*f*) et contient donc la boule $W_\varepsilon = \{y: |y| < \varepsilon\}$ incluse dans l'enveloppe convexe de deux boules mentionnées.

Pour la raison d'homothétie, il est clair que, quel que soit $\rho > 0$, on a l'inclusion $W_\rho \subset \overline{A(V_{N\rho/\varepsilon})}$. En particulier, on a $W_{\varepsilon/N} \subset \overline{A(V_1)}$, ce qu'il nous fallait.

Montrons à présent que l'ensemble $A(V_1)$ lui-même (et non seulement sa fermeture) contient la boule $W_{\varepsilon/(2N)}$. Soit $y \in W_{\varepsilon/(2N)}$.

Puisqu'on a démontré que $W_{\varepsilon/(2N)} \subset \overline{A(V_{1/2})}$, il est possible de choisir un point $y_1 \in A(V_{1/2})$ autant proche que l'on veut du point y. Par exemple, on peut le faire de façon à avoir $|y - y_1| < \varepsilon/(4N)$. Vu que $W_{\varepsilon/(4N)} \subset \overline{A(V_{1/4})}$, on peut trouver de même un point $y_2 \in A(V_{1/4})$ tel que $|y - y_1 - y_2| < \varepsilon/(8N)$. En continuant le procédé on construit, pour tout $n = 1, 2, \ldots$, un point $y_n \in A(V_{1/2^n})$ tel que $|y - y_1 - y_2 - \ldots - y_n| < \varepsilon/(2^{n+1}N)$.

Par construction on a $y = \sum_{n=1}^{\infty} y_n$. Or, $y_n = Ax_n$, où $x_n \in V_{1/2^n}$, de sorte que $|x_n| < 1/2^n$. L'espace X étant complet, la série $x_1 + x_2 + \ldots$ converge (12.37*b*) ; soit $x = \sum_{n=1}^{\infty} x_n$. Comme l'opérateur A est continu, $Ax = A\left(\sum_1^{\infty} x_n\right) = \sum_1^{\infty} Ax_n = \sum_1^{\infty} y_n = y$. De plus, $|x| \leqslant \sum_1^{\infty} |x_n| < \sum_1^{\infty} \frac{1}{2^n} = 1$. Par conséquent, la boule $W_{\varepsilon/(2N)}$ est contenue dans l'image de la boule V_1, ce qu'on affirmait.

Toujours pour la raison d'homothétie, on a $W_\rho \subset A(V_{\varepsilon\rho/(2N)})$ pour tout $\rho > 0$. En particulier, il résulte de $|x - x_0| < \delta$ que $|Ax - Ax_0| = |A(x - x_0)| < \delta\varepsilon/(2N)$, de sorte que l'image $A(U)$ de la boule $U = \{x : |x - x_0| < \delta\}$ contient la boule $\{y : |y - Ax_0| < \delta\varepsilon/(2N)\}$. Il en découle que l'image de tout ensemble ouvert $G \subset X$ est un ensemble ouvert dans Y, et le théorème est complètement démontré.

c. Conséquence. *Si* A *est une application continue et isomorphe* (12.14 *j*) *d'un espace normé complet* X *sur un espace normé complet* Y, *alors l'application inverse* A^{-1} *est elle aussi continue.*

Démonstration. Dans ce cas, l'opérateur inverse A^{-1} est défini d'une façon univoque et est évidemment linéaire de même que A. En vertu du théorème *b*, l'image réciproque par l'opérateur A^{-1} de tout ensemble ouvert $G \subset X$ est l'ensemble ouvert $AG \subset Y$. En particulier, l'image réciproque de la boule $\{x : |x| < \varepsilon\}$ contient une boule $\{y : |y| < \delta\}$, ce qui signifie la continuité de l'application A^{-1}.

d. Conséquence. *Si un espace vectoriel* L *est complet par rapport à chacune des deux normes* $|x|_1$ *et* $|x|_2$, *alors l'existence d'une constante* c_1 *telle que* $|x|_2 \geqslant c_1 |x|_1$ *pour tout* $x \in L$ *implique l'existence d'une constante* c_2 *telle que* $|x|_1 \geqslant c_2 |x|_2$ *pour tout* $x \in L$; *les normes* $|x|_1$ *et* $|x|_2$ *s'avèrent donc équivalentes* (12.35).

Démonstration. Considérons l'application identique A de l'espace normé X que l'on obtient en munissant L de la norme $|x|_2$ sur l'espace normé Y que l'on obtient en munissant L de la norme

$|x|_1$. En raison de l'inégalité $|x|_2 \geqslant c_1 |x|_1$, cette application est continue. Par hypothèse et selon c l'application inverse est continue elle aussi, d'où le résultat cherché (12.71*d*).

e. Supposons qu'un espace complet X soit mis sous forme de somme directe de deux sous-espaces fermés X_1 et X_2, de sorte que, pour tout vecteur $x \in X$, on a une représentation unique

$$x = x_1 + x_2, \quad x_1 \in X_1, \quad x_2 \in X_2.$$

L'opérateur P_1 qui à tout vecteur x fait correspondre sa composante x_1 s'appelle *projecteur* (ou *projection*) *sur le sous-espace* X_1 ; d'une manière analogue, l'opérateur P_2 faisant correspondre à tout vecteur x sa composante x_2 s'appelle *projecteur* (ou *projection*) *sur le sous-espace* X_2. Ces opérateurs sont évidemment linéaires, mais il n'est point évident qu'ils soient continus. Nous verrons que les *opérateurs* P_1 *et* P_2 *sont continus* en supposant l'espace X complet, les sous-espaces X_1 et X_2 fermés, et en utilisant le théorème sur l'application ouverte.

En plus de la norme initiale $|x| = |x|_1$, introduisons dans l'espace X la norme

$$|x|_2 = |x_1|_1 + |x_2|_1.$$

Il est évident que $|x|_2$ vérifie les axiomes de la norme. Nous avons aussi $|x|_1 \leqslant |x_1|_1 + |x_2|_1 = |x|_2$. Montrons que l'espace X est complet par rapport à la norme $|x|_2$. Soit $\{x^{(n)}\}$ une suite de Cauchy pour la norme $|x|_2$; il résulte de l'égalité $|x^{(n)} - x^{(m)}|_2 = |x_1^{(n)} - x_1^{(m)}|_1 + |x_2^{(n)} - x_2^{(m)}|_1$ que les suites $\{x_1^{(n)}\}$ et $\{x_2^{(n)}\}$ sont de Cauchy pour la norme $|x|_1$. L'espace X étant complet, il existe les limites $x_1 = \lim\limits_{n\to\infty} x_1^{(n)}$, $x_2 = \lim\limits_{n\to\infty} x_2^{(n)}$; les sous-espaces X_1 et X_2 étant fermés, on a $x_1 \in X_1$, $x_2 \in X_2$. Posons $x = x_1 + x_2$. Nous avons $|x - x^{(n)}|_2 = |x_1 - x_1^{(n)}|_1 + |x_2 - x_2^{(n)}|_1 \to 0$, de sorte que x est la limite de la suite $\{x^{(n)}\}$ pour la norme $|x|_2$, ce qui démontre que X est complet par rapport à la norme $|x|_2$. En appliquant *d* on voit que les normes $|x|_1$ et $|x|_2$ sont équivalentes ; en particulier, il existe une constante c telle que l'inégalité

$$|x|_2 = |x_1|_1 + |x_2|_1 \leqslant c\,|x|_1 = c\,|x|$$

a lieu pour tout $x \in X$; mais alors on a aussi

$$|P_1 x| = |x_1|_1 \leqslant c\,|x|, \quad |P_2 x| = |x_2|_1 \leqslant c\,|x|,$$

ce qui démontre la continuité des opérateurs P_1 et P_2.

f. Soient toujours X un espace complet, somme directe de deux sous-espaces fermés X_1 et X_2, et P_1 et P_2 les projecteurs correspondants. Soient A_1 un opérateur linéaire continu dans X_1 et A_2 un opérateur linéaire continu dans X_2. Définissons dans l'espace X

l'opérateur A d'après la formule

$$Ax = A(x_1 + x_2) = A_1x_1 + A_2x_2.$$

L'opérateur A est évidemment linéaire.

L'opérateur A *est continu dans l'espace* X.

En effet,

$$Ax = A_1x_1 + A_2x_2 = A_1P_1x + A_2P_2x$$

et, les opérateurs P_1 et P_2 étant bornés dans l'espace X d'après e, nous avons

$$|Ax| \leqslant ||A_1|| \cdot ||P_1|| \cdot |x| + ||A_2|| \cdot ||P_2|| \cdot |x| = c|x|,$$

ce qu'il nous fallait.

12.73. Convergence d'une suite d'opérateurs linéaires.

a. Dans 12.71*b* on a introduit la norme

$$||A|| = \sup_{|x| \leqslant 1} |Ax|$$

dans l'espace L (X, Y) des opérateurs linéaires d'un espace normé X dans un espace normé Y.

Une suite $A_1, A_2, \ldots$ d'opérateurs *converge* pour cette norme vers un opérateur A si, pour tout $\varepsilon > 0$, il existe un numéro N tel que l'inégalité

$$\sup_{|x| \leqslant 1} |Ax - A_nx| \leqslant \varepsilon$$

est remplie pour tout $n \geqslant N$.

b. Montrons que *l'espace* L (X, Y) *est complet si* Y *l'est*. Soit $A_1, A_2, \ldots$ une suite de Cauchy d'opérateurs linéaires de X dans Y, de sorte que pour tout $\varepsilon > 0$ il existe un numéro N tel que l'inégalité

$$||A_n - A_m|| \leqslant \varepsilon \qquad (1)$$

a lieu pour $n, m \geqslant N$. Quel que soit $x \in X$, nous avons d'après 12.71 (1):

$$|A_nx - A_mx| \leqslant ||A_n - A_m|| \, |x| \leqslant \varepsilon |x|,$$

de sorte que les vecteurs $A_nx \in Y$ forment une suite de Cauchy dans l'espace Y. L'espace Y étant complet, il existe un vecteur $y \in Y$ tel que $y = \lim_{n \to \infty} A_nx$; posons $y = Ax$ et montrons que A est un opérateur linéaire borné qui est la limite (dans l'espace L (X, Y)) de la suite A_n. L'égalité

$$A(\alpha x + \beta y) = \lim_{n \to \infty} A_n(\alpha x + \beta y) = \lim_{n \to \infty} (\alpha A_nx + \beta A_ny) =$$

$$= \alpha \lim_{n \to \infty} A_nx + \beta \lim_{n \to \infty} A_ny = \alpha Ax + \beta Ay$$

montre que A est linéaire. Ensuite, pour $|x| \leqslant 1$, on a

$$|Ax - A_m x| = |(A - A_m)x| = \lim_{n\to\infty} |(A_n - A_m)x| \leqslant \varepsilon \qquad (2)$$

en vertu de (1) et pour $m \geqslant N$; il en résulte que $A - A_m$, donc aussi A, est un opérateur borné. Enfin, l'inégalité (2) montre que, pour $m \geqslant N$, on a l'inégalité

$$\| A - A_m \| \leqslant \varepsilon,$$

de sorte que $A = \lim_{m\to\infty} A_m$ pour la norme de l'espace L (X, Y).

Si Y est l'axe réel R_1, l'espace L (X, Y) = L (X, R_1) est bien complet. Cet espace (celui de toutes les fonctionnelles linéaires continues sur l'espace X) est appelé *espace dual* (ou *dual* tout court) *de* X et désigné par X^*.

c. En particulier, l'espace L (X, X) des opérateurs linéaires bornés dans un espace de Banach X est complet. Dans la suite il est désigné par L (X).

d. Il arrive que les opérateurs A, A_1, A_2, ... possèdent la propriété $A_n x \to Ax$ pour tout $x \in X$, mais $\| A_n - A \|$ ne tend pas vers zéro. (Un exemple sera donné dans 12.75*g*.) Le lemme suivant nous sera utile:

L e m m e. *Si les normes des opérateurs* A_1, A_2, ... *sont majorées par une même constante c et que la relation limite* $Ax = \lim_{n\to\infty} A_n x$ *soit valable pour tous les éléments x d'un ensemble Q partout dense dans* X, *alors* $Ax = \lim_{n\to\infty} A_n x$ *quel que soit* $x \in X$.

D é m o n s t r a t i o n. Soit $x = \lim_{k\to\infty} x_k$, où $x_k \in Q$. Pour un $\varepsilon > 0$ donné, trouvons un numéro k de façon à avoir $|x - x_k| < \varepsilon/(3c)$, puis un numéro N tel que l'inégalité $|Ax_k - A_n x_k| < \varepsilon/3$ ait lieu pour $n \geqslant N$. Alors, pour $n \geqslant N$, on aura

$$|Ax - A_n x| \leqslant |Ax - Ax_k| + |Ax_k - A_n x_k| +$$

$$+ |A_n x_k - A_n x| \leqslant \|A\| |x - x_k| + \frac{\varepsilon}{3} +$$

$$+ \|A_n\| |x_k - x| \leqslant c\frac{\varepsilon}{3c} + \frac{\varepsilon}{3} + c\frac{\varepsilon}{3c} = \varepsilon,$$

ce qui veut dire que $Ax = \lim_{n\to\infty} A_n x$.

Une suite d'opérateurs A_1, A_2, ... possédant la propriété $A_n x \to Ax$ pour tout $x \in X$ sera appelée *fortement convergente vers l'opérateur* A, et ce dernier opérateur sera dit *limite forte de la suite* A_n.

12.74. P r i n c i p e d e l a b o r n e u n i f o r m e.

a. T h é o r è m e (B a n a c h et S t e i n h a u s). *Si les normes d'une suite d'opérateurs linéaires continus* A_1, A_2, ... *d'un espace de*

Banach X *dans un espace normé* Y *forment une suite non bornée*

$$\sup_n \|A_n\| = \infty,$$

alors il existe, dans toute boule $U_\rho(x_0) = \{x \in X : |x - x_0| < \rho\}$, *un point* x *tel que*

$$\sup_n |A_n(x)| = \infty.$$

Démonstration. La suite des fonctions $A_1(x)$, $A_2(x)$, ... dont chacune est bornée dans la boule $|x| \leqslant 1$ n'est pas bornée uniformément dans cette boule. Pour la raison d'homothétie, elle n'est uniformément bornée sur aucune boule $|x| \leqslant r$. Qui plus est, elle n'est bornée uniformément sur aucune boule de la forme $U_r(x_0) = \{x : |x - x_0| \leqslant r\}$ car, si les vecteurs $A_n(x)$ et $A_n(x_0)$ étaient bornés pour $x \in U_r(x_0)$, il en serait de même des vecteurs $A_n(x - x_0) = A_n(x) - A_n(x_0)$, ce qui est impossible puisque $x - x_0$ parcourt la boule de rayon r et de centre 0. Choisissons donc dans la boule $U_\rho(x_0)$ un élément x_1, $|x_1 - x_0| < \rho$ pour lequel la valeur de l'un quelconque des opérateurs A_n (désignons-le pour le moment par A_1) dépasse en norme l'unité :

$$|A_1(x_1)| > 1.$$

L'opérateur A_1 étant continu, il existe une boule $U_{\rho_1}(x_1)$ contenue entièrement dans la boule initiale $U_\rho(x_0)$ et dans laquelle l'inégalité

$$|A_1(x)| > 1$$

est satisfaite.

Trouvons à l'intérieur de cette boule un élément x_2 et un opérateur $A_2(x)$ de façon à avoir $|A_2(x)| > 2$, puis une autre boule $U_{\rho_2}(x_2)$ contenue dans la précédente en tout point de laquelle on a

$$|A_2(x)| > 2 \quad \left(\rho_2 < \frac{1}{2}\rho_1\right)$$

En continuant nous aboutissons à une suite des boules emboîtées de rayons $\rho_1, \rho_2, \ldots$ tendant vers zéro. Pour le point x commun à toutes ces boules (qui existe en vertu de la complécité de l'espace X et du lemme 3.74*d*), on a les inégalités

$$|A_1(x)| > 1, \quad |A_2(x)| > 2, \ldots, |A_n(x)| > n, \ldots,$$

ce qu'il fallait démontrer.

b. Conséquence. *Si* $A_1, A_2, \ldots$ *est une suite d'opérateurs linéaires continus d'un espace de Banach* X *dans un espace normé* Y *et si, pour tout vecteur* x *de l'espace de Banach* X, *la suite des vecteurs* $A_1x, A_2x, \ldots$ *est bornée, alors les normes des opérateurs* $A_1, A_2, \ldots$ *sont majorées par une même constante.*

c. Conséquence. *Si une suite* $A_1, A_2, \ldots$ *d'opérateurs linéaires continus d'un espace de Banach* X *dans un espace de Banach* Y *est telle que, pour tout* $x \in X$, *les vecteurs* $y_n = A_n x$ *ont une limite* $y \in Y$, *alors l'application* $A: \{x \to \lim A_n x\}$ *est un opérateur linéaire continu de* X *dans* Y.

Démonstration. Soient x_1 et x_2 deux vecteurs quelconques de l'espace X, α_1 et α_2 des constantes arbitraires. En passant à la limite pour $n \to \infty$ dans l'égalité $A_n(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A_n x_1 + \alpha_2 A_n x_2$ on obtient $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2$, de sorte que l'application A est linéaire. La suite des vecteurs $A_n x$ étant convergente donc bornée pour tout $x \in X$, on voit, d'après *b*, que les normes des opérateurs A_n sont bornées: $\| A_n \| \leqslant C$. Par conséquent, nous avons $| A_n x | \leqslant \| A_n \| \leqslant C$ pour tout x, $| x | \leqslant 1$, donc $| Ax | = \lim\limits_{n \to \infty} | A_n x | \leqslant C$; ainsi, l'opérateur A est borné dans la boule unité, donc continu, ce qu'il fallait démontrer.

De plus, la suite des opérateurs A_n converge fortement vers l'opérateur A (12.73*d*).

d. Le raisonnement précédent fournit l'estimation suivante pour la norme de l'opérateur A:

$$\| A \| \leqslant \sup_n \| A_n \|.$$

On peut la préciser. Soit $c = \varlimsup \| A_n \|$, et supposons que, pour un $\varepsilon > 0$ donné, on dégage une sous-suite A_{n_k} $(k = 1, 2, \ldots)$ pour laquelle $\| A_{n_k} \| \leqslant c + \varepsilon$. Comme, évidemment, $Ax = \lim\limits_{k \to \infty} A_{n_k} x$ pour tout $x \in X$, on a

$$\| A \| \leqslant \sup \| A_{n_k} \| \leqslant c + \varepsilon$$

d'après ce qui précède; vu que ε est arbitraire, on a

$$\| A \| \leqslant \varlimsup \| A_n \|.$$

Dans certains cas concrets, on peut mettre le symbole $<$ dans cette inégalité (on en verra un exemple dans 12.75*g*).

e. Lemme. *Supposons qu'une suite d'opérateurs* A_n *converge fortement vers un opérateur* A *et une suite de vecteurs* x_n *converge (en norme) vers un vecteur* x. *Alors* $Ax = \lim\limits_{n \to \infty} A_n x_n$.

Démonstration. En vertu du principe de la borne uniforme, les normes des opérateurs A_n sont bornées par une constante C. Donc

$$| Ax - A_n x_n | \leqslant | Ax - A_n x | + | A_n x - A_n x_n | \leqslant$$
$$\leqslant | (A - A_n) x | + C | x - x_n |.$$

Les deux termes dans le second membre convergent vers zéro pour $n \to \infty$, ce qui démontre le lemme.

f. Dans *a-e*, on peut remplacer la suite d'opérateurs A_n par une fonction A (t) à valeurs opératorielles, définie sur un ensemble $T = \{t\}$ et la convergence pour $n \to \infty$ par la convergence suivant une direction S définie sur l'ensemble T (4.12).

Dans les paragraphes qui suivent nous exposons certaines applications importantes du principe de la borne uniforme.

12.75. Espace des suites bornées et ses sous-espaces.

a. Désignons par X l'espace vectoriel de toutes les suites réelles bornées $x = (\xi_1, \xi_2, \ldots)$ avec les opérations usuelles (par coordonnées) et avec la norme définie par la formule

$$\|x\| = \sup_n |\xi_n|.$$

Les axiomes de l'espace vectoriel normé sont vérifiés d'une façon évidente. De plus, *l'espace* X *est complet*; on peut le démontrer directement ou bien citer le théorème sur la complécité de l'espace R^S (M) de toutes les fonctions réelles bornées et continues sur un espace métrique M (12.23*f*); dans le présent cas, cet espace métrique est l'ensemble des nombres naturels muni de la métrique usuelle de la droite numérique.

b. Soit une suite de nombres réels $f_1, f_2, \ldots, \sum_{n=1}^{\infty} |f_n| < \infty$. Alors, pour tout $x = (\xi_1, \ldots, \xi_n, \ldots) \in X$, l'expression

$$f(x) = \sum_{n=1}^{\infty} f_n \xi_n \tag{1}$$

est définie et l'inégalité

$$|f(x)| \leqslant \sup_n |\xi_n| \sum_{n=1}^{\infty} |f_n| \tag{2}$$

a lieu.

L'expression (1) représente, évidemment, une fonctionnelle linéaire sur l'espace X. L'inégalité (2) montre que cette fonctionnelle est bornée sur la boule unité de l'espace X, elle est donc continue; de plus, sa norme a pour majoration

$$\|f\| \leqslant \sum_{n=1}^{\infty} |f_n|. \tag{3}$$

Considérons la valeur de la fonctionnelle f pour le vecteur $x_0 = (\xi_1, \xi_2, \ldots)$, où $\xi_k = \operatorname{sgn} f_k$ ($k = 1, 2, \ldots$). Notons que le vecteur x_0 appartient à la boule unité de l'espace X. Nous avons

$$f(x_0) = \sum_{k=1}^{\infty} f_k \operatorname{sgn} f_k = \sum_{k=1}^{\infty} |f_k|,$$

d'où

$$\|f\| = \sup_{|x|\leqslant 1} |f(x)| \geqslant |f(x_0)| = \sum_{k=1}^{\infty} |f_k|. \tag{4}$$

En comparant les inégalités (3) et (4) nous voyons que

$$\|f\| = \sum_{n=1}^{\infty} |f_n|. \tag{5}$$

Les fonctionnelles de la forme (1) n'épuisent pas l'ensemble de toutes les fonctionnelles linéaires continues sur l'espace X. Tout de même, pour certains sous-espaces de l'espace X, la formule (1) donne la forme générale de la fonctionnelle linéaire continue. L'un des sous-espaces de ce genre est considéré dans *c*.

c. Désignons par X_0 l'ensemble de tous les éléments $x = (\xi_1, \xi_2, \ldots) \in X$ pour lesquels $\lim\limits_{n\to\infty} \xi_n = 0$. Il est évident que X_0 est un sous-espace dans l'espace X. Prouvons que ce sous-espace est fermé. Soit

$$x_m = \{\xi_n^{(m)}\} \in X_0 \quad (m = 1, 2, \ldots) \text{ et } x = \{\xi_n\} = \lim_{m\to\infty} x_m.$$

Pour un $\varepsilon > 0$ donné, trouvons un numéro m tel que $\|x_m - x\| = \sup\limits_n |\xi_n^{(m)} - \xi_n| < \varepsilon/2$. Puis trouvons un numéro p tel que pour $n \geqslant p$ on a l'inégalité $|\xi_n^{(m)}| < \varepsilon/2$. Alors, pour tout $n \geqslant p$, on a également $|\xi_n| \leqslant |\xi_n - \xi_n^{(m)}| + |\xi_n^{(m)}| < \varepsilon$, ce qui signifie que $\lim\limits_{n\to\infty} \xi_n = 0$.

L'ensemble X_0 étant fermé dans un espace complet X, l'espace normé X_0 est complet.

d. Posons à présent $e_n = (0, \ldots, 0, 1, 0, \ldots)$, où 1 occupe la n-ième place. Pour tout $x = (\xi_1, \xi_2, \ldots) \in X_0$, nous avons

$$\left\|x - \sum_{n=1}^{m} \xi_n e_n\right\| = \|(\xi_1, \ldots, \xi_m, \xi_{m+1}, \ldots) - (\xi_1, \ldots, \xi_m, 0, \ldots)\| =$$
$$= \|(0, \ldots, 0, \xi_{m+1}, \xi_{m+2}, \ldots)\|$$

et, si le nombre N est choisi pour un $\varepsilon > 0$ donné de façon à avoir $|\xi_m| < \varepsilon$ pour $m > N$, nous obtenons $\left\|x - \sum\limits_{1}^{m} \xi_n e_n\right\| < \varepsilon$; ainsi

$$x = \sum_{1}^{\infty} \xi_n e_n,$$

la série étant convergente par rapport à la norme de l'espace X_0. En particulier, nous voyons que l'ensemble des $x = (\xi_1, \xi_2, \ldots)$, pour lesquels toutes les coordonnées ξ_n à partir de l'une quelconque sont nulles, représente un ensemble partout dense dans X_0.

e. Supposons qu'une suite $\{f_k\}$ soit telle que la série $\sum_1^\infty f_k\xi_k$ converge quel que soit $x=\{\xi_n\}\in X_0$. Alors la série $\sum_1^\infty |f_k|$ converge elle aussi. En effet, considérons les fonctionnelles linéaires

$$\varphi_n(x)=\sum_1^n f_k\xi_k \quad (n=1, 2, \ldots).$$

Par hypothèse, les valeurs de ces fonctionnelles ont une limite pour $n\to\infty$ quel que soit $x\in X_0$. Alors, d'après 12.74*b*, les normes des fonctionnelles φ_n sont majorées par une même constante C. En appliquant (5) on a pour tout n l'inégalité

$$\|\varphi_n\|=\sum_1^n |f_k|\leqslant C,$$

d'où la convergence de la série $\sum_1^\infty |f_k|$.

f. Montrons à présent que *l'expression* (1) *fournit la forme générale d'une fonctionnelle linéaire continue sur l'espace* X_0. Soit $f(x)$ une fonctionnelle linéaire continue sur l'espace X_0. Posons $f(e_k)=f_k$ et formons la suite des fonctionnelles linéaires continues $\varphi_n(x)=\sum_{k=1}^n f_k\xi_k$ $(n=1, 2, \ldots)$. Puisque l'égalité $x=\sum_{k=1}^\infty \xi_k e_k$ a lieu pour tout $x\in X_0$ et la fonctionnelle f est continue, on a

$$f(x)=f\Big(\sum_{k=1}^\infty \xi_k e_k\Big)=\sum_{k=1}^\infty \xi_k f(e_k)=\sum_{k=1}^\infty \xi_k f_k$$

pour tout $x\in X_0$. Nous voyons que la fonctionnelle f agit selon la formule (1), la série $\sum_1^\infty |f_k|$ étant convergente en vertu de *e*. Notre proposition est donc démontrée.

Notons encore que la norme $\|f\|_0$ de la fonctionnelle f dans l'espace X_0 est égale à sa norme $\|f\|=\sum_1^\infty |f_k|$ dans X tout entier. En effet, il est évident que $\|f\|_0\leqslant\|f\|$. D'autre part, en appliquant la fonctionnelle f au vecteur $x_n=\{\operatorname{sgn} f_1, \ldots, \operatorname{sgn} f_n, 0, 0, \ldots\}\in X_0$ nous avons $f(x_n)=\sum_1^n f_k \operatorname{sgn} f_k=\sum_1^n |f_k|$, d'où $\|f\|_0\geqslant\sum_1^n |f_k|$ pour tout $n=1, 2, \ldots$, de sorte que $\|f\|_0\geqslant\sum_1^\infty |f_k|$. Par conséquent, $\|f\|_0=\sum_1^\infty |f_k|=\|f\|$, ce qu'on affirmait.

g. Considérons en particulier la fonctionnelle $g_k(x) = \xi_k$ (qui à tout vecteur x fait correspondre sa k-ième coordonnée). Elle si déduit de (1) en posant $f_k = 1$, $f_m = 0$ pour $m \neq k$. La norme de cette fonctionnelle vaut 1 quel que soit $k = 1, 2, \ldots$ De pluse pour tout $x = (\xi_1, \xi_2, \ldots) \in X_0$, on a

$$\lim_{k\to\infty} g_k(x) = \lim_{k\to\infty} \xi_k = 0.$$

Ainsi, la suite des fonctionnelles g_k converge fortement vers zéro pour $k \to \infty$ (12.73*d*) bien que leurs normes ne le fassent pas.

h. Désignons par X_1 l'ensemble de tous les éléments $x = (\xi_1, \xi_2, \ldots) \in X$ admettant la limite finie de la suite ξ_n pour $n \to \infty$. L'ensemble X_1 forme, évidemment, un sous-espace de l'espace X et contient le sous-espace X_0 et l'espace unidimensionnel $\{\lambda e\}$ des éléments de la forme $\lambda e = \{\lambda, \lambda, \lambda, \ldots\}$; il est évident que X_1 est la somme directe de deux sous-espaces mentionnés. Le sous-espace X_1 est fermé dans l'espace X ; on peut le démontrer directement ou bien citer le théorème 12.23*f* en remarquant que X_1 peut être considéré comme espace de toutes les fonctions bornées et continues sur l'espace métrique composé des nombres naturels 1, 2, ... et du symbole ∞, avec une métrique dans laquelle les nombres 1, 2, ... sont isolés et $\infty = \lim_{n\to\infty} n$ (cf. 3.35*f*).

i. Dans le sous-espace X_1 il y a déjà une fonctionnelle linéaire continue qui n'est pas de la forme (1), notamment

$$L(x) = \lim_{n\to\infty} \xi_n.$$

Si l'on pouvait la mettre sous la forme (1) avec certains nombres $f_1, f_2, \ldots$, de sorte que

$$L(x) = \sum_{k=1}^{\infty} f_k \xi_k = \lim_{n\to\infty} \xi_n,$$

alors, en posant $x = e_m$, on obtiendrait $L(e_m) = f_m = 0$ $(m = 1, 2, \ldots)$; mais dans ce cas, pour $x = e = (1, 1, \ldots)$, on aurait $L(e) = \sum_{k=1}^{\infty} f_k \cdot 1 = 0$ en contradiction avec $L(e) = \lim 1 = 1$.

Ainsi, la fonctionnelle $L(x)$ n'est pas une fonctionnelle du type (1). Or elle est la limite forte (12.73*d*) de fonctionnelles de la forme (1) sur l'espace X_1 : évidemment on a, pour tout $x \in X_1$,

$$L(x) = \lim_{k\to\infty} g_k(x).$$

12.76. Sommation de suites bornées.

a. Nous savons qu'il existe des suites bornées, mais non convergentes, de nombres réels. Proposons-nous d'étendre la notion de limite d'une suite convergente aux suites de nombres réels qui divergent

eu sens usuel du mot. Autrement dit, il s'agit de faire correspondre à toute suite $x = \{a_n\}$ d'un sous-espace fermé $X^* \subset X$ contenant le sous-espace X_1 des suites convergentes un nombre Lim a_n appelé *limite généralisée* et vérifiant les conditions naturelles suivantes:

1) Lim $(\alpha a_n + \beta b_n) = \alpha$Lim $a_n + \beta$Lim b_n quels que soient des suites $\{a_n\}$ et $\{b_n\}$ de X^* et des nombres réels α et β;

2) Lim $a_n = \lim a_n$ pour toute suite convergente a_n;

3) Lim a_n est une fonctionnelle bornée et continue sur X^*.

b. En tant qu'exemple, considérons la limite au sens de Cesàro. Par définition, la limite au sens de Cesàro d'une suite $\{a_n\}$, désignée par C-lim a_n, est égale à la limite ordinaire (si cette dernière existe) de la suite des moyennes arithmétiques des a_n:

$$\text{C-lim } a_n = \lim_{n\to\infty} \frac{a_1 + \dots + a_n}{n}.$$

On peut montrer que la limite de Cesàro vérifie les conditions 1)-3) sur son domaine d'existence; nous n'insistons pas là-dessus, parce que, un peu plus bas, un théorème plus général sera démontré. Notons que la limite de Cesàro peut bien exister dans certains cas où la limite ordinaire n'existe pas. Il est, par exemple, évident que

$$\text{C-lim } (0, 1, 0, 1, \dots) = \frac{1}{2}.$$

c. Les fonctionnelles de la forme 12.75 (1) ne conviennent pas pour la construction d'une limite généralisée, parce que même la limite ordinaire $\lim\limits_{n\to\infty} \xi_n$ sur le sous-espace X_1 ne peut être mise sous cette forme (12.75*i*). La limite ordinaire est, on l'a vu dans 12.75*i*, la limite forte des fonctionnelles spéciales de la forme 12.75 (1); il faut donc étudier la possibilité d'obtenir une limite généralisée comme limite forte de fonctionnelles 12.75 (1). Pour avoir une suite de fonctionnelles de cette forme, on doit se donner une matrice infinie $T = \| t_{km} \|$, $k, m = 1, 2, \dots$ dont les lignes définissent les fonctionnelles

$$T_k(x) = \sum_{m=1}^{\infty} t_{km}\xi_m.$$

S'il existe la limite ordinaire de la suite des nombres $T_k(x)$ pour $k \to \infty$, nous l'appelons T-limite de la suite $\{\xi_n\}$ et désignons

$$T(x) = \lim_{k\to\infty} T_k(x) = \text{T-}\lim_{n\to\infty} \xi_n.$$

Comment doit être la matrice T pour que le domaine de définition de la fonctionnelle $T(x)$ contienne toutes les suites convergentes et pour que les conditions 1)-3) caractérisant une limite généralisée soient satisfaites? La réponse est donnée par le théorème suivant.

Théorème (Toeplitz). *La fonctionnelle* $T(x)$ *est une limite généralisée si, et seulement si, les conditions suivantes sont vérifiées* :

$$\sum_{m=1}^{\infty} |t_{km}| \leqslant c, \text{ où } c \text{ ne dépend pas de } k; \tag{1}$$

$$\lim_{k\to\infty} \sum_{m=1}^{\infty} t_{km} = 1; \tag{2}$$

$$\lim_{k\to\infty} t_{km} = 0 \quad (m = 1, 2, \ldots). \tag{3}$$

Démonstration. Prouvons la nécessité des conditions (1)-(3). Si la fonctionnelle $T(x)$ est définie, il en est de même de toutes les fonctionnelles $T_k(x)$ pour n'importe quelle suite $x = \{\xi_n\}$ convergeant vers zéro. Il en résulte en vertu de 12.75*e* que chaque série $\sum_{m=1}^{\infty} |t_{km}|$ converge. Ensuite, d'après le théorème de Banach-Steinhaus 12.74*b*, il découle de la convergence de la suite $T_k(x)$, pour tout $x \in X_0$, que les normes des fonctionnelles T_k sont bornées; en écrivant ces normes comme dans 12.75*b* nous voyons que la condition (1) est vérifiée. En appliquant la T-limite à la suite $x = (1, 1, \ldots)$ nous voyons que la condition (2) est aussi vérifiée. En appliquant la T-limite au vecteur e_k (12.75*d*) nous établissons la condition (3).

Démontrons à présent que les conditions (1)-(3) sont suffisantes pour que la fonctionnelle $T(x)$ soit une limite généralisée.

Si les valeurs $T(x) = \lim_{k\to\infty} T_k(x)$ et $T(y) = \lim_{k\to\infty} T_k(y)$ sont définies pour certaines suites $x = \{\xi_n\}$ et $y = \{\eta_n\}$, alors le nombre $T(\alpha x + \beta y) = \lim T_k(\alpha x + \beta y) = \alpha \lim T_k(x) + \beta \lim T_k(y)$ est défini quelles que soient les constantes α et β; ainsi le domaine de définition X_T de la fonctionnelle $T(x)$ est un espace vectoriel, la fonctionnelle elle-même étant linéaire sur ce domaine. Pour la suite $e = (1, 1, 1, \ldots)$, nous avons:

$$T_k(e) = \sum_{m=1}^{\infty} t_{km}, \quad T(e) = \lim_{k\to\infty} T_k(e) = \lim_{k\to\infty} \sum_{m=1}^{\infty} t_{km} = 1$$

en vertu de la condition (2). Par conséquent, lorsqu'on vérifie l'égalité (2) figurant dans la définition d'une limite généralisée, il suffit de se borner aux éléments $x \in X_0$. En vertu de la condition (1), les fonctionnelles T_k dans l'espace X_0 sont bornées en norme par la constante c. Ensuite, pour les éléments $x = (\xi_1, \ldots, \xi_n, 0, 0, \ldots)$, nous avons $\lim T_k(x) = \lim_{k\to\infty} \sum_{1}^{n} t_{km}\xi_m = 0$ d'après la condition (3). Ces éléments formant un ensemble partout dense dans X_0 (12.75*d*), le résultat cherché découle du lemme 12.73*d*. Enfin, le fait que la

fonctionnelle T (x) est bornée sur le sous-espace X_T résulte de 12.74*c*, et 12.74*d* fournit la majoration de la norme

$$\| T \| \leqslant \varliminf_{k \to \infty} \| T_k \| = \varliminf_{k \to \infty} \sum_{n=1}^{\infty} | t_{kn} |.$$

Comme $Te = 1$, on a une minoration de la norme

$$\| T \| \geqslant 1.$$

Il reste à montrer que le sous-espace X_T est fermé dans l'espace X. Considérons la fermeture $\overline{X}_T$ du sous-espace X_T. L'ensemble X_T est partout dense dans $\overline{X}_T$, les fonctionnelles $T_k(x)$ convergent en tout point $x \in X_T$, leurs normes étant bornées; vu 12.73*d*, il en résulte que les fonctionnelles $T_k(x)$ convergent aussi sur $\overline{X}_T$. Nous voyons que le domaine de convergence X_T de la suite T_k contient sa fermeture $\overline{X}_T$; donc, $\overline{X}_T = X_T$ et le théorème est complètement démontré.

12.77. Exemples.

a. A la limite ordinaire $\lim \xi_n$ (qui n'est définie que sur X_1), il correspond la matrice

$$\begin{Vmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{Vmatrix}.$$

b. La limite de Cesàro (12.76*b*) est donnée par la matrice

$$\begin{Vmatrix} 1 & 0 & 0 & \dots \\ 1/2 & 1/2 & 0 & \dots \\ 1/3 & 1/3 & 1/3 & \dots \\ \dots & \dots & \dots & \dots \end{Vmatrix},$$

pour laquelle les hypothèses du théorème de Toeplitz sont remplies; la limite de Cesàro possède donc toutes les propriétés d'une limite généralisée.

c. En tant que troisième exemple nous signalons toute une classe de matrices qui contient les deux précédentes comme ses cas particuliers. Soient $p_0 > 0$, $p_1 \geqslant 0$, $p_2 \geqslant 0$, ... une suite et $P_n = \sum_{0}^{n} p_k$ $(n = 0, 1, 2, \dots)$; alors la matrice

$$\begin{Vmatrix} 1 & 0 & 0 & \dots & 0 & \dots \\ p_1/P_1 & p_0/P_1 & 0 & \dots & 0 & \dots \\ p_2/P_2 & p_1/P_2 & p_0/P_2 & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_n/P_n & p_{n-1}/P_n & p_{n-2}/P_n & \dots & p_0/P_n & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{Vmatrix}$$

définit une méthode de sommation due à Voronoï. Les conditions (1) et (2) du théorème de Toeplitz sont ici immédiates. La condition (3), pour $m = 1$, est équivalente à la condition $p_n/P_n \to 0$; or, puisque

$$\frac{p_{n-m}}{P_n} \leqslant \frac{p_{n-m}}{P_{n-m}},$$

(3) résulte de la condition $p_n/P_n \to 0$ pour tout m. Par conséquent, *la condition* $p_n/P_n \to 0$ *est nécessaire et suffisante pour que la matrice de Voronoï soit une matrice de Toeplitz.* Si $p_0 = 1$, $p_1 = p_2 = \ldots = 0$, on revient à la sommation ordinaire ; si $p_0 = p_1 = p_2 = \ldots = 1$, on retrouve la sommation au sens de Cesàro.

12.78. Signalons encore quelques propriétés des T-limites.

a. Pour certaines T-limites, l'espace X_T peut se confondre avec l'espace X_1, comme c'est le cas pour la limite ordinaire $\lim_{n\to\infty} \xi_n$. On se demande comment séparer ces cas « peu intéressants» de limite généralisée. Il y a un théorème de A. Broudno [2], d'après lequel $X_T = X_1$ si, et seulement si, il existe une constante δ_0 telle que, pour tout $x = \{\xi_n\} \in X$, on a l'inégalité

$$\overline{\lim_{n\to\infty}} |T_n(x)| \geqslant \delta_0 \overline{\lim_{n\to\infty}} |\xi_n|.$$

Pourtant, cette condition est assez dure à vérifier. Il y a encore une condition de l'égalité $X_T = X_1$ en termes des nombres t_{kn} eux-mêmes, mais elle n'est que suffisante (Agnew) : $X_T = X_1$ si l'inégalité

$$\lim_{n\to\infty} \sum_n \Big[t_{nn} - \sum_{k\neq n} |t_{kn}|\Big] > 0$$

a lieu [18].

b. D'autre part, est-il possible de construire une matrice T pour laquelle $X_T = X$? Cela s'avère impossible (cf. exercice 8). Tout de même, il résulte de certaines considérations d'ordre général qu'il existe une limite généralisée $\operatorname{Lim} \xi_n$ définie sur tout l'espace X et telle que $\operatorname{Lim}_{n\to\infty} \xi_n = \operatorname{Lim}_{n\to\infty} \xi_{n+1}$ [20] ; cependant, une telle limite généralisée ne peut être donnée par une formule explicite.

c. Pour certaines matrices T, la quantité $T(x)$ peut sortir de l'intervalle $\Delta x = [\lim \xi_n, \overline{\lim}\, \xi_n]$ contenant toutes les valeurs d'adhérence de la suite ξ_n. Ceci a lieu, par exemple, pour la matrice

$$\begin{Vmatrix} 2 & -1 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 2 & -1 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & 2 & -1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{Vmatrix}$$

et pour l'élément $x = (1, 0, 1, 0, 1, 0, \ldots)$. Il est naturel de poser la question suivante : quelles sont les conditions à imposer à

la matrice T pour que toutes les valeurs d'adhérence de la suite $T_1(x)$, $T_2(x)$, ... (pour tout $x \in X$) appartiennent à l'intervalle $\Delta(x)$? La réponse est donnée par le théorème suivant de Robinson (cf. exercice 9): la propriété demandée a lieu si, et seulement si, $\lim_{n\to\infty} \| T_n \| = 1$.

§ 12.8. Algèbres normées

12.81.a. Un espace normé U qui est en même temps une algèbre (12.18*a*) s'appelle *algèbre normée* si $x_n \to x$ (par rapport à la norme de U) implique $x_n y \to xy$ et $yx_n \to yx$ pour tout $y \in U$.

b, Ainsi, l'ensemble L (X) de tous les opérateurs bornés agissant dans un espace de Banach X est un espace normé complet (12.73*b*) et en même temps une algèbre (12.19*i*, 12.71*j*). Dans cette algèbre, la norme satisfait à l'inégalité 12.71 (2)

$$\| AB \| \leqslant \| A \| \| B \|; \tag{1}$$

il en résulte que L (X) est une algèbre normée: notamment, si $A_n \to A$ et si B est un opérateur quelconque de L (X), alors on a

$$\| A_n B - AB \| = \| (A_n - A) B \| \leqslant \| A_n - A \| \| B \| \to 0$$

de sorte que $A_n B \to AB$.

c. Il se trouve qu'une inégalité du type (1) est valable pour n'importe quelle algèbre normée complète, après le passage à une autre norme (on le verra dans 12.88). C'est pourquoi on peut remplacer la condition de continuité de la multiplication sous la forme « $x_n \to x$ implique $x_n y \to xy$ et $yx_n \to yx$ pour tout y » par une condition plus forte

$$| xy | \leqslant | x | | y | \tag{2}$$

quels que soient x et y de U.

d. Dans ce qui suit nous supposons, en plus de l'axiome (2), qu'une algèbre normée considérée possède une unité e (12.18*c*) et que $| e | = 1$. (La dernière supposition est satisfaite automatiquement dans l'algèbre des opérateurs linéaires agissant dans un espace normé X, l'unité étant l'opérateur identique.)

12.82.a. L'unité d'une algèbre normée, comme de n'importe quelle algèbre, est un élément inversible car $ee = e$. Montrons que *dans une algèbre normée complète* U, *toute la boule* $\{x : | e - x | < 1\}$ *est formée des éléments inversibles.*

Pour le démontrer, considérons la série

$$y = e + (e - x) + (e - x)^2 + \ldots \tag{1}$$

D'après la condition (2), on a $| (e - x)^n | \leqslant | e - x |^n$, donc la série converge en vertu du critère de Weierstrass 12.37*c*. En la

multipliant par $x = e - (e - x)$, nous avons

$$y\,[e - (e - x)] =$$

$$= [e + (e - x) + (e - x)^2 + \ldots] - [(e - x) + (e - x)^2 + \ldots] = e.$$

Par conséquent, la somme de la série (3) est justement l'élément inverse de x.

b. Il résulte de l'estimation

$$|e - y| = |(e - x) + (e - x)^2 + \ldots| \leqslant \frac{|e - x|}{1 - |e - x|}$$

que $x \to e$ implique $y \to e$. On peut dire que l'opérateur (non linéaire) de multiplication par $y = x^{-1}$ *est continu pour* $x = e$.

12.83.a. Désignons par O l'ensemble de tous les éléments inversibles d'une algèbre normée complète U. Montrons que O *est un ensemble ouvert dans* U *et l'opérateur* x^{-1} *est continu sur tout le* O.

Etant donné que $xx^{-1} = e$, on a, en vertu de (2) pour tout h tel que $|h| < 1/|x^{-1}|$: $|(x + h)x^{-1} - e| \quad |hx^{-1}| \leqslant |h|\,|x^{-1}| < 1$, de sorte que l'élément $(x + h)x^{-1}$ est inversible d'après 12.82*a*, i.e. qu'il existe un élément $z(h)$ tel que $(x + h)x^{-1}z(h) = e$. Mais alors $x + h$ est aussi inversible: son inverse est évidemment l'élément $x^{-1}z(h)$. Si $h \to 0$, alors $(x + h)x^{-1} \to xx^{-1} = e$, de sorte que $z(h) \to e$ en vertu de 12.82*b*; il en découle $(x + h)^{-1} = x^{-1}z(h) \to x^{-1}$, ce qui démontre la continuité de l'opérateur x^{-1} sur l'ensemble O.

b. Nous avons vu qu'un élément inversible x appartient à l'ensemble O avec une boule de rayon $r \geqslant 1/|x^{-1}|$. Ceci veut dire que $|x^{-1}| \geqslant 1/r$; ainsi, lorsque x s'approche de la frontière de l'ensemble O, nous avons, bien entendu, $r \to 0$, et la norme de l'élément x^{-1} augmente indéfiniment.

12.84. Un élément non inversible z se trouvant sur la frontière du domaine O est un « diviseur généralisé de zéro »; ceci veut dire qu'il existe une suite d'éléments $y_1, y_2, \ldots$ telle que $|y_n| \geqslant c > 0$, mais $zy_n \to 0$. A savoir, nous posons $y_n = x_n^{-1}/|x_n^{-1}|$, où $x_n \in O$ et $x_n \to z$. Alors $|zy_n| \leqslant |(z - x_n)y_n| + |x_n y_n| \leqslant |z - x_n| + 1/|x_n^{-1}| \to 0$ (12.83*b*), ce qu'il nous fallait.

En général, tout diviseur généralisé de zéro z n'est pas inversible, car, pour un z inversible, $zy_n \to 0$ implique $z^{-1}zy_n = y_n \to 0$. Mais un élément non inversible qui n'est pas une limite d'éléments inversibles peut bien ne pas être un diviseur généralisé de zéro (exercice 10).

12.85.a. Une algèbre complexe normée et complète U sera appelée par la suite *algèbre de Gelfand*.

Quel que soit un élément x d'une algèbre de Gelfand, l'expression $e - \lambda x$ est un élément inversible pour n'importe quels λ com-

plexes suffisamment petits, par exemple pour $|\lambda| < 1/|x|$ si $x \neq 0$; donc, on a d'après 12.82*a* :

$$(e - \lambda x)^{-1} = e + \lambda x + \lambda^2 x^2 + \dots \tag{1}$$

Le rayon de convergence exact de la série (1) vaut (12.66)

$$\rho = \frac{1}{\overline{\lim_{n\to\infty}} \sqrt[n]{|x^n|}} ; \tag{2}$$

dans le cas où $\rho = \infty$, la série converge dans tout le plan des λ.

L'élément $x - \mu e$ est inversible pour tout $|\mu|$ suffisamment grand, par exemple pour $|\mu| > |x|$; cela résulte directement de la formule $x - \mu e = -\mu(e - \mu^{-1}x)$. L'ensemble de tous les μ pour lesquels l'élément $x - \mu e$ n'est pas inversible s'appelle *spectre de l'élément x*. La fonction $(x - \mu e)^{-1}$ est définie sur le complémentaire du spectre. D'après 12.83*a*, ce complémentaire est un ensemble G ouvert dans le plan des μ, le spectre étant fermé. Ensuite, il résulte de 12.83*a* que $(x - \mu e)^{-1}$ est une fonction continue de μ (à valeurs dans U) dans le domaine G. Montrons que, de plus, elle est une fonction analytique (12.66) sur G. On a l'égalité

$$\left[\frac{(x-(\mu+h)e)^{-1} - (x-\mu e)^{-1}}{h}\right](x-(\mu+h)e)(x-\mu e) =$$

$$= \frac{(x-\mu e) - (x-(\mu+h)e)}{h} = e \tag{3}$$

qui montre que l'élément entre crochets est inversible ; son inverse $(x - (\mu + h)e)(x - \mu e)$ a la limite $(x - \mu e)^2$ lorsque $h \to 0$, d'où l'existence de la limite

$$\lim_{h\to 0} \frac{(x-(\mu+h)e)^{-1} - (x-\mu e)^{-1}}{h} = [(x-\mu e)^2]^{-1}. \tag{4}$$

Cela veut dire que $(x - \mu e)^{-1}$ est une fonction analytique dans le domaine G, ce qu'il nous fallait.

b. T h é o r è m e. *Le spectre de tout élément x d'une algèbre de Gelfand* U *n'est pas vide.*

D é m o n s t r a t i o n. Soit Γ une circonférence dans le plan des μ de centre au point 0 et de rayon $r > |x|$. Considérons l'intégrale

$$I = \frac{1}{2\pi i} \oint_{\Gamma} (x - \mu e)^{-1} \, d\mu \tag{5}$$

(qui existe en vertu de la continuité de la fonction $(x - \mu e)^{-1}$ sur la ligne Γ). Calculons-la en faisant la substitution $\mu^{-1} = \lambda$ et en utili-

sant le développement (1) et les formules 10.33*a*. Nous avons :

$$I = -\frac{1}{2\pi i}\oint_{|\mu|=r} \mu^{-1}(e-\mu^{-1}x)^{-1}\,d\mu = \frac{1}{2\pi i}\oint_{|\lambda|=1/r} (e-\lambda x)^{-1}\frac{d\lambda}{\lambda} =$$

$$= \frac{1}{2\pi i}\oint_{|\lambda|=1/r}\sum_{m=0}^{\infty}\lambda^m x^m\frac{d\lambda}{\lambda} = \frac{1}{2\pi i}\sum_{m=0}^{\infty}x^m\oint_{|\lambda|=1/r}\lambda^{m-1}\,d\lambda = -e.$$

Or, si le spectre de l'élément x était vide, la fonction $(x-\mu e)^{-1}$ serait analytique dans tout le plan des μ, et l'intégrale (5) serait nulle d'après le théorème de Cauchy. Le théorème est démontré.

c. Conséquence (théorème de Gelfand-Mazur). *Si une algèbre de Gelfand* U *est un corps, i.e. si tout son élément x non nul admet un inverse, alors l'algèbre* U *est le corps des nombres complexes.*

En effet, soient x un élément quelconque de l'algèbre U et μ un nombre de son spectre, de sorte que l'élément $x-\mu e$ n'ait pas d'inverse. Or, par hypothèse, le seul élément qui n'ait pas d'inverse est 0 ; donc $x-\mu e=0$, $x=\mu e$, ce qu'il fallait démontrer.

12.86. a. Considérons le cas où l'algèbre U est l'algèbre U (C_n) de tous les opérateurs linéaires dans un espace C_n de dimension finie doté d'une norme quelconque (nous avons vu dans 12.35*e* que chaque norme dans C_n est équivalente à n'importe quelle autre). Tout opérateur linéaire A est dans le présent cas continu, car les coordonnées du vecteur Ax sont des fonctions linéaires, donc continues, des coordonnées du vecteur x. Ainsi, l'algèbre U (C_n) se confond avec l'algèbre L (C_n) de tous les opérateurs bornés dans l'espace C_n.

Le spectre, au sens de la définition 12.85*a*, de l'élément A est dans ce cas l'ensemble de toutes les valeurs propres de l'opérateur A : en effet, l'opérateur $A-\mu E$ n'est pas inversible si, et seulement si, $\det\|A-\mu E\|=0$; or, c'est l'équation définissant les valeurs propres de l'opérateur A. Nous voyons que la définition du spectre dans 12.85*a* est équivalente, pour le cas en question, à la définition du spectre d'un opérateur A donnée dans 12.19*d*. Le spectre d'un opérateur A dans 12.19*d*, avec ses multiplicités, nous a permis de mettre d'une façon isomorphe l'algèbre de tous les opérateurs P (A) sous la forme de l'algèbre des corpus sur le spectre de l'opérateur A ; il a aussi permis de construire un épimorphisme de l'algèbre U (G) des fonctions analytiques dans un domaine G contenant le spectre sur l'algèbre P (A).

b. Il s'avère que les morphismes analogues existent dans le cas général de n'importe quelle algèbre de Gelfand. Soit $S=S_x$ le spectre d'un élément $x\in$ U ; nous savons que S est un ensemble non vide fermé et borné dans le plan complexe. Soit U (S) l'algèbre de toutes les fonctions analytiques $f(\lambda)$ sur l'ensemble S (chacune d'elles est analytique dans un domaine contenant l'ensemble S).

Théorème. *Il existe un morphisme de l'algèbre* U (S) *dans l'algèbre* U *qui transforme la fonction* $f(\lambda) \equiv 1$ *en* e, *la fonction* $f(\lambda) \equiv \lambda$ *en l'élément* x *et toute suite de fonctions* $f_n(\lambda)$ ($n = 1, 2, \ldots$) *qui converge vers une fonction* $f(\lambda)$ *uniformément sur un domaine* $H \supset S$ *en une suite d'éléments* $f_n \in$ U *convergeant en norme vers l'élément* f *qui correspond à la fonction* $f(\lambda)$.

Démonstration. Le morphisme cherché est donné par la formule

$$f = \frac{1}{2\pi i} \oint_{\Gamma} (\lambda e - x)^{-1} f(\lambda)\, d\lambda, \tag{1}$$

où Γ est un contour fermé se trouvant dans le domaine d'analyticité de la fonction $f(\lambda)$ et contournant (une fois) l'ensemble S. En vertu du théorème de Cauchy, l'intégrale (1) ne dépend pas du choix de ce contour. Le fait que la fonction $f(\lambda) \equiv 1$ devient, par l'application (1), l'élément e est prouvé dans 12.85*b* ; on démontre de façon analogue que la fonction $f(\lambda) \equiv \lambda$ devient l'élément x. Il est évident que la formule (1) définit une application linéaire de U (S) dans U ; il nous faut montrer que le produit de deux fonctions $f(\lambda)$ et $g(\lambda)$ devient produit des éléments correspondants f et g.

Nous partons de l'égalité

$$(\lambda e - x)^{-1} (\mu e - x)^{-1} = \frac{(\lambda e - x)^{-1} - (\mu e - x)^{-1}}{\mu - \lambda} \tag{2}$$

qui se déduit de 12.85 (3) en remplaçant $\mu + h$ par λ. Considérons deux courbes fermées Γ_f et Γ_g contournant l'ensemble S dans le domaine d'analyticité commun aux fonctions $f(\lambda)$ et $g(\lambda)$, de sorte que la courbe Γ_g enveloppe la courbe Γ_f sans avoir avec elle de points communs. En intégrant l'égalité (2), multipliée préalablement par $f(\lambda)\, g(\mu)/(2\pi i)^2$, d'abord le long de la courbe Γ_f, puis le long de Γ_g et en permutant les intégrations d'après 10.24, nous avons

$$\begin{aligned}
&\frac{1}{2\pi i} \oint_{\Gamma_f} (\lambda e - x)^{-1} f(\lambda)\, d\lambda \frac{1}{2\pi i} \oint_{\Gamma_g} (\mu e - x)^{-1} g(\mu)\, d\mu = \\
&= \frac{1}{2\pi i} \oint_{\Gamma_f} (\lambda e - x)^{-1} f(\lambda) \left\{ \frac{1}{2\pi i} \oint_{\Gamma_g} \frac{g(\mu)\, d\mu}{\mu - \lambda} \right\} d\lambda + \\
&+ \frac{1}{2\pi i} \oint_{\Gamma_g} (\mu e - x)^{-1} g(\mu) \left\{ \frac{1}{2\pi i} \oint_{\Gamma_f} \frac{f(\lambda)\, d\lambda}{\lambda - \mu} \right\} d\mu.
\end{aligned}$$

Comme λ est un point intérieur du domaine limité par la courbe Γ_g, la première intégrale entre accolades vaut $g(\lambda)$; comme μ est à l'extérieur du domaine limité par la courbe Γ_f, la deuxième intégrale

entre accolades est nulle. Définitivement, nous avons l'égalité

$$\frac{1}{2\pi i}\oint_{\Gamma_f}(\lambda e-x)^{-1}f(\lambda)\,d\lambda\,\frac{1}{2\pi i}\oint_{\Gamma_g}(\mu e-x)^{-1}g(\mu)\,d\mu=$$
$$=\frac{1}{2\pi i}\oint_{\Gamma_f}(\lambda e-x)^{-1}f(\lambda)\,g(\lambda)\,d\lambda$$

qui montre que la formule (1) fait correspondre au produit de deux fonctions $f(\lambda)$ et $g(\lambda)$ le produit des éléments correspondants f et g.

Examinons la dernière affirmation du théorème. Supposons qu'une suite de fonctions $f_n(\lambda)$ converge vers une fonction $f(\lambda)$ uniformément dans un domaine G contenant l'ensemble S. Choisissons un contour fermé Γ dans le domaine G; nous aurons l'estimation

$$\|f-f_n\|=\left\|\frac{1}{2\pi i}\oint_{\Gamma}(\lambda e-x)^{-1}[f(\lambda)-f_n(\lambda)]\,d\lambda\right\|\leqslant$$
$$\leqslant\sup_{\lambda\in\Gamma}|f(\lambda)-f_n(\lambda)|\,\frac{1}{2\pi i}\oint_{\Gamma}\|(\lambda e-x)^{-1}\|\,|d\lambda|,$$

d'où $\lim\limits_{n\to\infty}\|f-f_n\|=0$. Le théorème est démontré.

Notons que l'application (1) n'est en général pas un monomorphisme et peut transformer une fonction $f(\lambda)\neq 0$ en 1 élément nul de l'algèbre U.

c. En particulier, pour tout élément $x\in U$, sont définies les fonctions e^{tx}, $\cos tx$, $\sin tx$; les propriétés d'un morphisme impliquent l'égalité

$$e^{(t_1+t_2)x}=e^{t_1x}e^{t_2x}\quad(t_1,\,t_2\in C).$$

Il résulte de la dernière affirmation du théorème que ces fonctions peuvent également être définies à l'aide de séries de puissances:

$$e^{tx}=\sum_{0}^{\infty}\frac{t^nx^n}{n!},$$

$$\cos tx=1-\frac{t^2}{2!}x^2+\frac{t^4}{4!}x^4-\ldots,$$

$$\sin tx=tx-\frac{t^3}{3!}x^3+\frac{t^5}{5!}x^5-\ldots$$

12.87. On peut caractériser le spectre de tout élément de la forme $f(x)$:

Théorème. *Si $f(\lambda)$ est une fonction analytique sur le spectre S_x d'un élément $x\in U$, alors le spectre $S_{f(x)}$ de l'élément $f(x)$ (12.86b) se confond avec l'ensemble des valeurs de $f(\lambda)$ pour $\lambda\in S_x$.*

Démonstration. Soient $\lambda_0\in S_x$, $\mu_0=f(\lambda_0)$. La fonction analytique $f(\lambda)-\mu_0$ s'annule pour $\lambda=\lambda_0$, donc admet la

représentation

$$f(\lambda) - \mu_0 = (\lambda - \lambda_0)\, g(\lambda),$$

où $g(\lambda)$ est encore une fonction analytique dans le même domaine que $f(\lambda)$. En vertu des propriétés d'un morphisme f, nous avons

$$f(x) - \mu_0 e = (x - \lambda_0 e)\, g(x).$$

Or, si $f(x) - \mu_0 e$ est inversible, $x - \lambda_0 e$ l'est aussi (avec $g(x)\,(f(x) - \mu_0 e)^{-1}$ pour inverse), ce qui contredit la supposition. Donc $\mu_0 \in S_{f(x)}$. Réciproquement, soit $\mu_0 \in S_{f(x)}$; il existe un $\lambda_0 \in S_x$ tel que $f(\lambda_0) = \mu_0$. En effet, si la fonction $f(\lambda) - \mu_0$ ne s'annulait pas sur S_x, alors la fonction $g(\lambda) = 1/[f(\lambda) - \mu_0]$ serait analytique sur l'ensemble S_x, et l'élément correspondant $g(x) \in \mathrm{U}$ serait l'inverse de $f(x) - \mu_0 e$, ce qui contredirait la supposition $\mu_0 \in S_{f(x)}$. Le théorème est démontré.

12.88. Revenons au problème de changement de norme d'une algèbre U avec la multiplication continue (i.e. telle que $x_n \to x$ implique $x_n y \to xy$ et $yx_n \to yx$) ayant pour but de satisfaire à la condition 12.81(2).

Théorème. *Pour toute algèbre normée complète* U *possédant une unité, avec* $|\ x\ |_1$ *pour norme, il existe une norme équivalente* $|\ x\ |_2$ *pour laquelle* $|\ e\ |_2 = 1$, $|\ xy\ |_2 \leqslant |\ x\ |_2\, |\ y\ |_2$ *quels que soient* x *et* y *de* U.

Démonstration. Tout élément x de l'algèbre U engendre l'opérateur A_x de multiplication par x selon la formule $A_x y = xy$. Il résulte de l'hypothèse et des propriétés d'une algèbre que A_x est un opérateur linéaire continu. Dans l'algèbre L (U) de tous les opérateurs linéaires continus agissant dans U, les opérateurs de la forme A_x forment une sous-algèbre V à laquelle l'opérateur unitaire $E = A_e$ sert d'unité.

En raison de l'associativité de la multiplication, nous avons $A_x(yz) = x(yz) = (xy)z = (A_x y)z$. Il est aisé de voir que cette propriété caractérise les opérateurs de la sous-algèbre V. En effet, si un opérateur A est tel que, quels que soient y et z de U, on a l'égalité $A(yz) = Ay \cdot z$, alors en posant $Ae = x$ nous avons $Ay = A(ey) = (Ae)y = xy$, c'est-à-dire que A est l'opérateur de multiplication par x.

Ayant établie cette propriété montrons que *la sous-algèbre* V *est fermée dans l'algèbre* L (U). Supposons que les opérateurs $A_1, A_2, \ldots$ de V convergent (par rapport à la norme de L (U)) vers un opérateur A. Alors $A_n x$ converge vers Ax pour tout $x \in \mathrm{U}$. La multiplication étant continue, nous avons $A(xy) = \lim A_n(xy) = \lim (A_n x \cdot y) = \lim A_n x \cdot y = Ax \cdot y$, d'où, d'après ce qui précède, $A \in V$.

Comme l'algèbre L (U) est complète (12.73*b*), la sous-algèbre $V \subset \mathrm{L}(\mathrm{U})$ fermée dans L (U) est complète elle aussi en tant qu'espace normé, avec la norme de L (U). A présent nous avons deux normes dans l'algèbre U : $|\ x_1\ |$ et

$$|\ x\ |_2 = \|\ A_x\ \| = \sup_{|y|_1 \leqslant 1} |\ A_x y\ |_1 = \sup_{|y|_1 \leqslant 1} |\ xy\ |_1$$

l'algèbre U étant complète par rapport à chacune d'elles. Ensuite, nous avons

$$|e|_2 = \|A_e\| = \|E\| = 1, \quad |x|_2 = \sup_{|y|_1 \leqslant 1} |xy|_1 \geqslant \left| x \frac{e}{|e|_1} \right|_1 = \frac{|x|_1}{|e|_1},$$

d'où

$$|x|_2 \geqslant c_1 |x|_1, \quad c_1 = \frac{1}{|e|_1}.$$

D'après 12.72*d*, les normes $|x|_1$ et $|x|_2$ sont équivalentes, ce qu'il fallait démontrer.

§ 12.9. Propriétés spectrales des opérateurs linéaires

12.91. Un opérateur linéaire borné A agissant dans un espace de Banach X appartient à l'algèbre L (X) de tous les opérateurs linéaires bornés agissant dans l'espace X. En tant qu'élément de cette algèbre il possède un spectre S_A (12.85*a*) qui est l ensemble des nombres complexes λ pour lesquels $A - \lambda E$ n'a pas d'opérateur inverse borné. Dans le cas de dimension finie ($X = C_n$), le spectre de l'opérateur A se réduit, comme nous l'avons vu dans 12.86*a*, à un nombre fini, par exemple m, de points distincts qui sont les valeurs propres de l'opérateur A. Nous savons que l'espace C_n est alors décomposable en somme directe de m sous-espaces invariants dans chacun desquels l'opérateur A a un seul point pour spectre et admet une description complète (12.17*f*). En cas de dimension infinie, le spectre de l'opérateur A est un ensemble compact non vide du plan, inclus dans le cercle $|\lambda| \leqslant \|A\|$ ou bien, plus exactement, dans le cercle $|\lambda| \leqslant \overline{\lim_{n\to\infty}} \sqrt[n]{\|A^n\|}$ (12.85*a*) ; dans le cas général, on n'y peut rien ajouter (cf. exercice 11 où l'on construit un opérateur ayant pour spectre un ensemble compact arbitraire du plan).

12.92.a. En cas de dimension infinie, un nombre λ appartenant au spectre d'un opérateur A n'est pas forcément une valeur propre de l'opérateur A. On peut même dire que, dans ce cas, c'est non plus la notion de valeur propre qui devient naturelle, mais celle de *valeur propre généralisée* : on appelle ainsi un nombre λ qui admet une suite de vecteurs $x_1, x_2, \ldots$ telle que $|x_n| \geqslant c > 0$ et $Ax_n - \lambda x_n \to 0$. Toute valeur propre d'un opérateur en est évidemment une valeur propre généralisée. *Toute valeur propre généralisée d'un opérateur* A *appartient à son spectre* ; en effet, si l'on a $Ax_n - \lambda x_n \to 0$ pour une suite x_n et que l'opérateur $A - \lambda E$ admette un inverse borné, alors $x_n = (A - \lambda E)^{-1}(A - \lambda E)\, x_n \to 0$.

b. Montrons à présent que *tout point sur la frontière du spectre d'un opérateur* A *est sa valeur propre généralisée.* Soit λ un point sur la frontière du spectre ; comme $A - \lambda E$ est la limite des opérateurs inversibles $A - \mu E$, où $\mu \notin S_A$, l'opérateur $A - \lambda E$ est un diviseur généralisé de zéro en vertu de 12.84, donc il existe une suite d'opéra-

teurs P_n telle que $\| P_n \| \geqslant c > 0$, mais $(A - \lambda E) P_n \to 0$ dans l'algèbre L (X). Pour tout opérateur P_n, nous choisissons un vecteur y_n tel que $| y_n | = 1$, $| P_n y_n | \geqslant c/2$. En posant $x_n = P_n y_n$, nous avons $| x_n | \geqslant c/2$, $| (A - \lambda E) x_n | = | (A - \lambda E) P_n y_n | \leqslant \leqslant \| (A - \lambda E) P_n \| | y_n | \to 0$, ce qu'il nous fallait.

Quant aux points intérieurs du spectre d'un opérateur A, ils ne sont pas forcément des valeurs propres généralisées (exercice 10).

12.93. Le théorème suivant rend parfois plus simple l'étude d'un opérateur:

Théorème. *Supposons que le spectre S_A d'un opérateur A soit la réunion de deux ensembles fermés disjoints S_1 et S_2. Alors l'espace X est décomposable en somme directe de deux sous-espaces fermés X_1 et X_2 qui sont invariants par A, de sorte que le spectre de A considéré sur le sous-espace X_1 est l'ensemble S_1 et celui considéré sur X_2 est S_2.*

Démonstration. Nous utilisons le morphisme de l'algèbre U (S_A) des fonctions $f(\lambda)$ analytiques sur S_A dans l'algèbre L (X) établi dans 12.86*b*. Ce morphisme est réalisé par la formule le 12.86(1)

$$f \to \frac{1}{2\pi i} \oint_\Gamma (\lambda E - A)^{-1} f(\lambda)\, d\lambda,$$

où Γ est une courbe fermée contournant l'ensemble S_A dans le domaine d'analyticité de la fonction $f(\lambda)$. Dans le cas où l ensemble S_A est réunion de ses parties fermées deux à deux disjointes, il peut en être de même de la courbe Γ. Dans le présent cas, l'ensemble S_A est réunion de deux ensembles fermés S_1 et S_2 sans points communs, et la courbe Γ peut se composer de deux courbes fermées Γ_1 et Γ_2, la première contournant l'ensemble S_1, la deuxième S_2.

La fonction $e_1(\lambda)$ qui vaut 1 sur l'ensemble S_1 et 0 sur S_2 appartient à l'algèbre U (S_A); l algèbre U (S_A) contient également la fonction $e_2(\lambda)$ valant 0 sur l'ensemble S_1 et 1 sur S_2. Ces fonctions possèdent les propriétés évidentes:

$$e_1(\lambda) + e_2(\lambda) \equiv 1 \text{ (sur } S_A), \quad e_1^2(\lambda) = e_1(\lambda),$$
$$e_2^2(\lambda) = e_2(\lambda), \quad e_1(\lambda)\, e_2(\lambda) = e_2(\lambda)\, e_1(\lambda) = 0.$$

Désignons par E_1 et E_2 les opérateurs linéaires correspondant respectivement aux fonctions $e_1(\lambda)$ et $e_2(\lambda)$. Vu les propriétés d'un morphisme, nous avons

$$E_1 + E_2 = E, \quad E_1^2 = E_1, \; E_2^2 = E_2, \; E_1 E_2 = E_2 E_1 = 0.$$

Soient X_1 l'ensemble des solutions (dans l'espace X) de l'équation $E_1 x = x$ et X_2 celui des solutions de l'équation $E_2 x = x$. En particulier, tout vecteur de la forme $x = E_1 y$, pour n'importe quel

$y \in X$, est solution de l'équation $E_1 x = x$ car $E_1 (E_1 y) = E_1^2 y = E_1 y$. Il est évident que X_1 et X_2 sont des sous-espaces dans l'espace X et sont fermés en vertu de la continuité des opérateurs E_1 et E_2. Si $z \in X_1 \cap X_2$, alors $z = E_1 z = E_2 z$; or $E_1 z = E_1 (E_1 z) = E_1 (E_2 z) = 0$, d'où $z = E_1 z = 0$. Ainsi, l'intersection des sous-espaces X_1 et X_2 ne contient que le vecteur nul. En appliquant l'opérateur $E = E_1 + E_2$ à un vecteur y quelconque on trouve $y = E_1 y + E_2 y$, le premier terme appartenant à X_1, le deuxième à X_2. L'espace X est donc décomposé en somme directe des sous-espaces X_1 et X_2.

Soit $x \in X_1$. Alors on a $Ax = A (E_1 x) = E_1 (Ax)$, donc Ax appartient aussi au sous-espace X_1; par conséquent, X_1 est invariant par l'opérateur A. D'une façon analogue, X_2 est invariant par A.

Il reste à démontrer la conclusion du théorème. Posons $A_1 = AE_1$; A_1 se confond avec A sur le sous-espace X_1 et est nul sur X_2. D'autre part, on peut écrire

$$A_1 = \frac{1}{2\pi i} \oint_{\Gamma} \lambda e_1 (\lambda) (\lambda E - A)^{-1} \, d\lambda.$$

Ici l'on peut remplacer la courbe Γ par Γ_1 car la fonction $e_1 (\lambda)$ est nulle sur la courbe Γ_2. Ceci fait, on peut remplacer la fonction $e_1 (\lambda)$ par 1; définitivement, on aura

$$A_1 = \frac{1}{2\pi i} \oint_{\Gamma_1} \lambda (\lambda E - A)^{-1} \, d\lambda.$$

Ensuite, quel que soit μ,

$$A_1 - \mu E_1 = \frac{1}{2\pi i} \oint_{\Gamma_1} (\lambda - \mu) (\lambda E - A)^{-1} \, d\lambda.$$

Supposons que μ soit à l'extérieur de la courbe Γ_1; construisons l'opérateur

$$Q_\mu = \frac{1}{2\pi i} \oint_{\Gamma_1} \frac{(\lambda E - A)^{-1} \, d\lambda}{\lambda - \mu}.$$

Les mêmes raisonnements que dans 12,86*b* impliquent

$$(A_1 - \mu E_1) Q_\mu = \frac{1}{2\pi i} \oint_{\Gamma_1} (\lambda - \mu) \frac{(\lambda E - A)^{-1} \, d\lambda}{\lambda - \mu} =$$

$$= \frac{1}{2\pi i} \oint_{\Gamma_1} (\lambda E - A)^{-1} \, d\lambda = E_1.$$

De la sorte, *l'opérateur* $A - \mu E$ *est inversible dans le sous-espace* X_1. Par conséquent, le spectre de l'opérateur A dans le sous-espace X_1 ne peut contenir que les points de l'ensemble S_1. De même, le spectre de A dans X_2 ne peut contenir que les points de S_2. Montrons que le spectre de A dans X_1 contient *tous* les points de l'ensemble S_1.

Soit $\lambda_0 \in S_1$. On a vu que l'opérateur $A - \lambda_0 E$ est inversible dans le sous-espace X_2, de sorte qu'il existe un opérateur Q_2 tel que $(A - \lambda_0 E)\, Q_2 y = y$ pour tout $y \in X_2$; si l opérateur $A - \lambda_0 E$ était inversible dans le sous-espace X_1, il existerait un opérateur Q_1 tel que $(A - \lambda_0 E)\, Q_1 x = x$ pour tout $x \in X_1$. On construirait alors un opérateur Q se confondant avec Q_1 dans X_1 et avec Q_2 dans X_2 et linéairement prolongé sur tout le X. En vertu de 12.72*f*, il serait un opérateur linéaire borné dans l'espace X. Evidemment, il serait en même temps l'inverse de $A - \lambda_0 E$. Or c'est impossible parce que $\lambda_0 \in S_A$. Le théorème est démontré.

12.94. Opérateurs compacts. Parmi tous les opérateurs agissant dans un espace normé, on peut dégager une classe importante d'opérateurs dont les propriétés sont les plus proches de celles des opérateurs dans un espace de dimension finie.

a. Définition. Un opérateur linéaire A agissant dans un espace normé X est dit *compact* s'il transforme tout ensemble borné $Q \subset X$ en un ensemble précompact (3.93*a*).

b. Dans un espace de dimension finie, tout opérateur linéaire est compact.

c. L'opérateur de Fredholm (12.98) sert d'exemple d'un opérateur compact dans l'espace $C^a(a, b)$.

d. L'opérateur identique E dans un espace de dimension infinie n'est pas compact, car il transforme la boule unité en elle-même, i.e. en un ensemble qui n'est pas précompact (12.36*b*).

12.95. Opérations sur les opérateurs compacts.

a. *La somme* $A_1 + A_2$ *de deux opérateurs compacts* A_1 *et* A_2 *est un opérateur compact.*

En effet, soient $Q \in X$ un ensemble borné et $\{x_n\}$ une suite de points de Q. L'opérateur A_1 étant compact, on peut extraire de la suite $\{x_n\}$ une sous-suite $\{x_n'\}$ de façon que $\{A_1 x_n'\}$ soit une suite de Cauchy, puis on peut extraire une sous-suite encore plus raréfiée $\{x_n''\}$ de façon que $\{A_2 x_n''\}$ soit une suite de Cauchy; alors $\{(A_1 + A_2) x_n''\}$ s'avère, évidemment, une suite de Cauchy, ce qu'il nous fallait.

b. *Le produit d'un opérateur compact* A *par n'importe quel opérateur borné* B (*l'ordre des opérateurs n'ayant pas d'importance*) *est un opérateur compact.*

En effet, soit $Q \in X$ un ensemble borné; alors BQ est borné également, donc ABQ est précompact; ainsi, l opérateur AB est compact. D'autre part, l'opérateur B transforme toute suite de Caucby en une suite de Cauchy et, par conséquent, l ensemble précompact AQ en un ensemble précompact; l'opérateur BA est donc compact lui aussi.

c. *En particulier, si l'opérateur compact* A *est inversible, l'espace* X *est de dimension finie.*

En effet, l'opérateur $E = AA^{-1}$ est compact d'après *b*, et l'on peut appliquer 12.94*d*.

d. *Si, pour tout* $n = 1, 2, \ldots$, *on a un opérateur compact* A_n *et si un opérateur* A *est tel que* $\lim \| A - A_n \| = 0$, *alors* A *est un opérateur compact.*

En effet, pour un $\varepsilon > 0$ donné, l'ensemble A_nQ (où Q est un ensemble borné quelconque inclus dans une boule $|x| \leqslant r$ et $\| A - A_n \| < \varepsilon$) est un ensemble précompact qui représente un εr-réseau pour l'ensemble AQ. Il en résulte que AQ est précompact (3.95) et l'opérateur A est compact.

12.96. Spectre d'un opérateur compact.

a. Lemme. *Pour un opérateur compact dans un espace de Banach* X, *toute valeur propre généralisée non nulle est une valeur propre ordinaire.*

Démonstration. Soit λ une valeur propre généralisée de l'opérateur compact A, de sorte qu'il existe une suite de vecteurs $x_1, x_2, \ldots$, $|x_n| \geqslant C > 0$ telle que $(A - \lambda E)x_n = q_n \to 0$. En vertu de la précompacité de l'ensemble $\{Ax_n\}$, il existe une suite de numéros $n_1, n_2, \ldots$ telle que les vecteurs Ax_{n_k} ont une limite dans l'espace X; posons $z = \lim\limits_{k\to\infty} Ax_{n_k}$. Alors la suite $\lambda x_{n_k} = Ax_{n_k} - q_{n_k}$ tend elle aussi vers z; en particulier, $|z| = \lim |\lambda x_{n_k}| \geqslant |\lambda| C > 0$. Ensuite, vu que $\lambda \neq 0$, on a $\lim\limits_{k\to\infty} x_{n_k} = z/\lambda$, $z = \lim Ax_{n_k} = (1/\lambda)Az$; on aboutit à l'égalité $Az = \lambda z$ qui démontre le lemme.

b. Lemme. *A l'extérieur de tout cercle* $|\lambda| \leqslant c$, $c > 0$, *un opérateur compact* A *ne peut avoir qu'un nombre fini de valeurs propres distinctes.*

Démonstration. Soient $\lambda_1, \lambda_2, \ldots$ des valeurs propres distinctes de l'opérateur A, avec $|\lambda_n| \geqslant c$. Soient $e_1, e_2, \ldots$ les vecteurs propres respectifs: $Ae_n = \lambda_n e_n$, $n = 1, 2, \ldots$. Les vecteurs propres associés à des valeurs propres distinctes sont linéairement indépendants (12.15*b*); l'enveloppe linéaire L_{n-1} des vecteurs $e_1, \ldots, e_{n-1}$ est donc un sous-espace propre de l'enveloppe linéaire L_n des vecteurs $e_1, \ldots, e_n$. D'après le lemme 12.36*a*, il existe un vecteur $h_n \in L_n$ tel que $|h_n| = 1$ et $|h_n - x| > 1/2$ pour tout $x \in L_{n-1}$. On peut noter $h_n = x_0 + \alpha e_n$, où $x_0 \in L_{n-1}$. Alors on a

$$Ah_n = A(x_0 + \alpha e_n) = Ax_0 + \alpha\lambda_n e_n = Ax_0 + \lambda_n(h_n - x_0) =$$
$$= (Ax_0 - \lambda_n x_0) + \lambda_n h_n.$$

Etant donné que $Ah_{n-1} + \lambda_n x_0 - Ax_0 \in L_{n-1}$, on en tire

$$|Ah_n - Ah_{n-1}| = |(Ax_0 - \lambda_n x_0 - Ah_{n-1}) + \lambda_n h_n| =$$
$$= |\lambda_n| \left| h_n - \frac{1}{|\lambda_n|}(Ah_{n-1} + \lambda_n x_0 - Ax_0) \right| \geqslant |\lambda_n| \cdot \frac{1}{2},$$

et l'on voit qu'il est impossible d'extraire une sous-suite convergente de la suite Ah_n. Ceci contredit la compacité de l'opérateur A. Le lemme est démontré.

e. Lemme. *A l'extérieur de tout cercle* $|\lambda| \leqslant c$, $c > 0$, *un opérateur compact* A *dans un espace de Banach* X *ne peut avoir qu'un nombre fini de points du spectre, ces points étant des valeurs propres de l'opérateur* A.

Démonstration. Tout point frontière du spectre de l'opérateur A est une valeur propre généralisée (12.92*b*), donc, d'après le lemme *a*, une valeur propre ordinaire ; par conséquent, il résulte du lemme *b* que, à l extérieur du cercle $|\lambda| \leqslant c$, l'opérateur compact A n'a qu'un nombre fini de points sur la frontière du spectre. Désignons ces points par $\lambda_1, \ldots, \lambda_n$; d'après *a*, ce sont des valeurs propres de l'opérateur A. Nous affirmons qu'ils épuisent tous les points du spectre de A se trouvant à l'extérieur du cercle $|\lambda| \leqslant c$. S'il y avait dans le spectre encore un point λ_0, $|\lambda_0| \geqslant c$, on mènerait par lui une droite allant vers l'infini et ne passant pas par le cercle $|\lambda| \leqslant c$ ni par les points $\lambda_1, \ldots, \lambda_n$; lo dernier point du spectre sur cette droite appartiendrait à la frontière du spectre sans coïncider avec aucun des points $\lambda_1, \ldots, \lambda_n$. La contradiction obtenue démontre le lemme.

d. Comme l'extérieur de tout cercle $|\lambda| \leqslant c$ ne contient, d'après le lemme *c*, qu'un nombre fini de points du spectre d'un opérateur compact, il est possible de numéroter tous les points du spectre par ordre de décroissance de leurs modules. Nous voyons que *le spectre d'un opérateur compact dans un espace de Banach représente un ensemble au plus dénombrable des valeurs propres isolées, avec* 0 *pour unique point limite.* Le point 0, pour un opérateur compact dans un espace de dimension infinie, est toujours un point du spectre (12.95*c*) (pas forcément valeur propre) ; les autres points du spectre forment un ensemble dénombrable, ou bien fini, ou enfin vide. Dans le dernier cas, on a l'égalité $\overline{\lim_{n\to\infty}} \sqrt[n]{\|A^n\|} = 0$ (12.85*a*), et l'opérateur A est nilpotent généralisé. L'analogue d'un tel opérateur dans un espaco de dimension finie est un opérateur nilpotent pour lequel $A^m = 0$ avec un m quelconque. Dans un espace de dimension finie, la structure d'un opérateur nilpotent peut être complètement caractérisée (dans une certaine base, il est donné par une matrice jordanienne dont tous les éléments diagonaux sont nuls). Dans le cas de dimension infinie, la structure d'un opérateur nilpotent généralisé n'est pas étudiée complètement [1].

12.97. Décomposition spectrale d'un opérateur compact.

a. Soit $\lambda \neq 0$ un point du spectre S_A d'un opérateur compact A ; comme il est isolé en vertu de 12.96*c*, on peut appliquer le théorème

12.93. L'espace X est alors décomposable en somme directe de deux sous-espaces fermés P_λ et Q_λ invariants par l'opérateur A, de sorte que, dans le premier d'eux, l'opérateur A a le nombre λ pour unique point du spectre et, dans le second, son spectre s'obtient en enlevant le point λ à S_A. Dans chacun des sous-espaces P_λ, Q_λ, l'opérateur A reste, bien entendu, compact; le point 0 n'appartenant plus à son spectre dans P_λ, il est inversible dans P_λ. Ceci veut dire que le sous-espace P_λ est *de dimension finie* (12.95*c*). Par conséquent, *tout point* $\lambda \neq 0$ *du spectre de l'opérateur* A *définit un sous-espace invariant de dimension finie*; *dans cet espace, naturellement, la structure de l'opérateur* A *est caractérisée par les moyens connus.*

b. On déduit de *a* une propriété importante d'un opérateur compact que voici :

Alternative de Fredholm. *Pour un* μ *complexe donné, deux possibilités peuvent se présenter: ou bien l'équation* $(E - \mu A)x = y$ *possède une solution unique par rapport à* x *pour tout* $y \in X$, *ou bien l'équation homogène* $(E - \mu A)\, x = 0$ *admet une solution non nulle.*

Démonstration. Il est évident que pour $\mu = 0$ c'est la première possibilité qui se réalise. Soit donc $\mu \neq 0$ et $\lambda = 1/\mu$; l'équation $(E - \mu A)\, x = y$ est alors équivalente à l'équation $(A - \lambda E)\, x = -\lambda y$. Si λ n'appartient pas au spectre de l'opérateur A, alors $A - \lambda E$ est inversible, et c'est le premier cas qui a lieu ; si λ appartient au spectre de A, alors λ est une valeur propre de A car $\lambda \neq 0$ (12.96*c*), et on a le deuxième cas.

Ainsi, *l'alternative de Fredholm est équivalente au fait que, pour l'opérateur* A, *tout nombre* $\lambda \in S_A$ *non nul est une valeur propre.* Nous avons vu que c'est bien la propriété des opérateurs compacts ; mais il existe une plus large classe d'opérateurs possédant cette propriété (par exemple, les opérateurs dont une puissance quelconque est compacte ; cf. exercice 13).

12.98. Opérateur intégral de Fredholm. Soit $q(s, t)$ une fonction complexe continue de deux variables réelles s et t variant dans un même intervalle $[a, b]$. L'intégrale

$$y(t) = \int_a^b q(s, t)\, x(s)\, ds, \tag{1}$$

pour toute fonction $x(t)$ continue sur $[a, b]$, représente une fonction définie toujours sur $[a, b]$ et continue en vertu de 9.81. Il est évident que la formule (1) définit un opérateur linéaire $y = Ax$ agissant dans l'espace $C^1[a, b]$ de toutes les fonctions complexes continues sur $[a, b]$, avec $\| x \| = \sup | x(t) |$ pour norme (12.39*b*) ; il s'appelle *opérateur de Fredholm*. Il résulte de l'inégalité

$$|y(t)| \leqslant \sup |x(s)| \int_a^b |q(s, t)|\, ds$$

que l opérateur A est borné et que sa norme ne dépasse pas le nombre

$$\sup_t \int_a^b |q(s, t)|\, ds.$$

Montrons que l'opérateur A est compact. Supposons que la fonction $x(t)$ parcourt un ensemble borné $Q \in C^0[a, b]$, de sorte que, par exemple, $\| x \| = \sup |x(t)| \leqslant r$.

Alors, pour $|t' - t''| \leqslant \delta$, on a

$$y(t') - y(t'') | \leqslant \sup |x(t)| \int_a^b |q(s, t') - q(s, t'')|\, ds \leqslant$$

$$\leqslant \sup |x(t)| \cdot \omega_q(\delta)\,(b - a),$$

où

$$\omega_q(\delta) = \sup_{|t'-t''|\leqslant\delta} |q(s, t') - q(s, t'')|.$$

On en déduit l'estimation suivante pour l'oscillation $\omega_y(\delta)$ de la fonction $y(t)$:

$$\omega_y(\delta) \equiv \sup_{|t'-t''|\leqslant\delta} |y(t') - y(t'')| \leqslant r\omega_q(\delta)\,(b - a)$$

cette estimation ne dépendant pas du choix de la fonction $x(t) \in Q$; comme la fonction $q(s, t)$ est continue, on a $\lim_{\delta\to 0} \omega_q(\delta) = 0$, donc aussi $\lim_{\delta\to 0} \omega_y(\delta) = 0$. Ainsi, l'ensemble $AQ \subset C^1(a, b)$ est uniformément borné et équicontinu ; d'après le théorème d'Arzelà (12.24*c*) l'ensemble AQ est précompact pour tout $Q \in C^1(a, b)$ borné ; l'opérateur A est donc compact, ce qu'il nous fallait.

En conséquence, nous voyons que toutes les propositions 12.96-12.97 sont valables pour l'opérateur de Fredholm. En particulier, a lieu l'alternative de Fredholm (12.97*b*) qui, dans le présent cas, prend la forme suivante :

Pour un μ complexe donné, deux possibilités peuvent se présenter : ou bien l'équation

$$x(t) - \mu \int_a^b q(s, t)\, x(s)\, ds = y(t)$$

possède une solution unique par rapport à $x(t)$ pour toute fonction $y(t) \in C^0(a, b)$, ou bien l'équation homogène

$$x(t) - \mu \int_a^b q(s, t)\, x(s)\, de = 0$$

admet une solution non nulle.

12.99. Opérateur intégral de Volterra. Soit toujours $q(s, t)$ une fonction continue des variables s et t parcourant l'intervalle $[a, b]$. L'intégrale

$$z(t) = [Vx](t) = \int_a^t q(s, t)\, x(s)\, ds \qquad (1)$$

diffère de l'intégrale 12.98 (1) par le fait que la limite constante b est remplacée par la limite variable t. De même que dans 12.98 (1) la fonction $y(t)$ est définie et continue dans $[a, b]$ (9.86a). L'opérateur linéaire V donné par la formule (1) s'appelle *opérateur de Volterra*. L'opérateur de Voltorra, de même que celui de Fredholm, est compact; on le démontre par la même méthode, en précisant un peu les estimations. Mais, contrairement à l'opérateur de Fredholm, *le spectre de l'opérateur de Volterra ne peut avoir de points non nuls* (c'est-à-dire, nous l'avons vu, de valeurs propres). Supposons le contraire: pour un $\lambda \neq 0$, il existe une fonction $x_0(t) \in C^a(a, b)$ telle que

$$Vx_0(t) \equiv \int_a^t q(s, t)\, x_0(s)\, ds = \lambda x_0(t). \qquad (2)$$

En posant $t = a$ on trouve $\lambda x_0(a) = 0$, donc $x_0(a) = 0$. Sans restreindre la généralité, on peut admettre que la fonction $x_0(s)$ ne s'annule identiquement dans aucun voisinage du point a de l'intervalle $[a, b]$, sinon on le ferait passer, sans que l'intégrale change de valeur, au point le plus proche de l'intervalle $]a, b]$ qui possède déjà la propriété mentionnée. Par conséquent, la fonction $m(\delta) = \sup_{a \leqslant t \leqslant a+\delta} |x(t)|$ est non nulle pour $\delta > 0$ et tend vers 0 pour $\delta \to 0$. Pour tout $\delta > 0$, on peut indiquer un point $t_\delta \in [a, a+\delta]$ tel que $|x_0(t_\delta)| = m(\delta)$. Maintenant, on obtient de (2) l'estimation suivante:

$$|\lambda x_0(t_\delta)| = |\lambda|\, m(\delta) \leqslant \max_{a \leqslant s \leqslant a+\delta} |x_0(s)| \cdot \int_a^{t_\delta} |q(s, t)|\, ds \leqslant c\delta m(\delta),$$

où

$$c = \sup_{t,\, s} |q(s, t)|.$$

En divisant par $m(\delta)$ on a

$$|\lambda| \leqslant c\delta$$

pour tout $\delta > 0$. Cette inégalité contredit la supposition $\lambda \neq 0$. La proposition est donc démontrée.

En appliquant 12.97b et 12.66 on voit que, pour tout $y(t)$, il existe une solution unique de l'équation de Voltorra

$$[E - \mu V]\, x(t) \equiv x(t) - \mu \int_a^t q(s, t)\, x(s)\, ds = y(t)$$

qui est représentée par la série

$$x(t) = (E - \mu V)^{-1} y(t) = \\ = y(t) + \mu V y(t) + \mu^2 V^2 y(t) + \ldots + \mu^n V^n y(t) + \ldots$$

On peut montrer que l'opérateur V^n, pour tout n, est encore un opérateur de Volterra de noyau $q_n(s, t)$ que l'on peut calculer par récurrence d'après les formules

$$q_1(s, t) = q(s, t),$$

$$q_n(s, t) = \int_s^t q_{n-1}(s, \sigma)\, q_1(\sigma, t)\, d\sigma$$

(cf., par exemple, [15]; dans le même livre on trouvera les exemples d'application des équations intégrales à la physique mathématique).

Exercices

1. Considérer trois espaces de fonctions sur la droite:
a) de toutes les fonctions continues et bornées;
b) de toutes les fonctions continues possédant la propriété $\lim_{|x| \to \infty} f(x) = 0$;
c) de toutes les fonctions continues dont chacune est nulle à l'extérieur d'un intervalle.
On munit ces espaces de la métrique

$$\rho(f, g) = \sup_t |f(x) - g(x)|.$$

Est-ce que les espaces mentionnés sont complets?

2. Indiquer dans l'espace $R^1(0, \infty)$ (de toutes les fonctions continues et bornées sur la demi-droite $0 < x < \infty$ avec $\|x(t)\| = \sup_t |x(t)|$ pour norme) un ensemble ayant la puissance du continu des fonctions $x_\alpha(t)$ telles que $\|x_\alpha(t)\| = 1$, $\|x_\alpha(t) - x_\beta(t)\| \geqslant 1$ pour $\alpha \neq \beta$.

Remarque. Il en résulte qu'il n'existe dans l'espace $R^1(0, \infty)$ aucun ensemble dénombrable partout dense.

3. Prouver que la fonctionnelle

$$F(y) = \int_0^{1/2} y(x)\, dx - \int_{1/2}^1 y(x)\, dx$$

est continue dans l'espace $R^1(0, 1)$; montrer que la borne supérieure de ses valeurs sur la boule unité fermée de l'espace $R^1(0, 1)$ vaut 1, cette borne n'étant atteinte sur aucun élément de la boule unité.

4. On sait que le lemme sur le parallélogramme (12.43*a*) est valable pour n'importe quels deux vecteurs x, y d'un certain espace normé X. Démontrer que la norme dans X est engendrée par le produit scalaire

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

5. Soit P l'algèbre de tous les polynômes $p(z)$ à coefficients complexes dans le cercle $Q = \{z: |z| \leqslant 1\}$, avec la norme $\|p(z)\| = \max_Q |p(z)|$. Cette algèbre contient 1 et sépare n'importe quels deux points du compact Q, mais le théorème de Stone 12.52*c* n'est pas valable pour elle, et l'algèbre P n'est pas dense dans l'algèbre $C^1(Q)$ de toutes les fonctions complexes continues dans le cercle Q.

6. Dans une algèbre normée $R^s(Q)$ (Q est un espace métrique), l'ensemble $I(F)$ de toutes les fonctions $f(x) \in R^s(Q)$ qui s'annulent identiquement sur un ensemble fermé $F \subset Q$ est un idéal fermé dans $R^s(Q)$. Montrer que si Q est un compact, alors tout idéal fermé $I \subset R^s(Q)$ coïncide avec l'idéal $I(F)$ pour un $F \subset Q$.

7. Supposons qu'une famille équicontinue $E \subset P^s(Q)$ de fonctions (P est un espace métrique, Q un compact) soit telle que, pour tout $t \in Q$, les valeurs des fonctions $x(t) \in E$ appartiennent à un précompact $P_t \subset P$. Démontrer qu'il existe un précompact $P_0 \subset P$ qui contient les valeurs de toutes les fonctions $x(t) \in E$ en tous les points $t \in Q$.

8. On se donne une matrice $T = \| t_{km} \|$ vérifiant l'hypothèse du théorème de Toeplitz 12.76c. Construire avec les nombres 1 et -1 une suite ξ_m n'ayant pas de T-limite.

9. Démontrer que la condition $\lim\limits_{n\to\infty} \| T_n \| = 1$ est nécessaire et suffisante pour que l'intervalle $[\underline{\lim}\, T_n(x), \overline{\lim}\, T_n(x)]$ (12.76) soit contenu dans l'intervalle $[\underline{\lim}\, x, \overline{\lim}\, x]$ quelle que soit la suite bornée $x = (\xi_1, \xi_2, \ldots)$.

10. Soit $\overline{C} = C^s(Q)$ l'algèbre de toutes les fonctions complexes $f(z)$ continues sur la circonférence $|z| = 1$ (avec la norme usuelle) et soit Z l'algèbre des fonctions $\varphi(z)$ analytiques dans le cercle $|z| < 1$ et continues dans le cercle $|z| \leqslant 1$, avec la même norme $\| \varphi \| = \sup |\varphi(z)|$. Montrer que

a) L'application qui fait correspondre à toute fonction $\varphi(z) \in Z$ la fonction limite $\varphi(e^{it}) \in C$ est un monomorphisme de Z dans C; par conséquent, on peut dire que l'algèbre Z est une sous-algèbre de l'algèbre C.

b) Z est une sous-algèbre fermée dans C.

c) Le spectre de l'opérateur A de multiplication par z dans l'espace C est la circonférence $|z| = 1$; le spectre du même opérateur dans l'espace Z est le cercle $|z| \leqslant 1$; de plus, les valeurs $|z| = 1$, et elles seules, sont les valeurs propres généralisées de l'opérateur A dans Z.

d) L'élément z est inversible dans l'algèbre C, non inversible dans l'algèbre Z et n'est pas un diviseur généralisé de zéro dans Z.

11. Soit Q un ensemble compact dans le plan des z et soit $C = C^s(Q)$ l'espace de toutes les fonctions complexes continues sur l'ensemble Q. Montrer que l'opérateur de multiplication par z a l'ensemble Q pour spectre.

12. On sait que, pour un opérateur A dans un espace de Banach X et pour un polynôme $p(\lambda)$, l'opérateur $p(A)$ est compact. Démontrer que tous les points du spectre de l'opérateur A (à l'exception possible des racines du polynôme $p(\lambda)$) sont ses valeurs propres.

13. Montrer que l'alternative de Fredholm a lieu pour un opérateur A dont une puissance quelconque est compacte.

14. Soient $p \geqslant 1$, $q \geqslant 1$ et $\frac{1}{p} + \frac{1}{q} = 1$. Pour deux fonctions quelconques $x(t)$ et $y(t)$ continues par morceaux sur un intervalle $a \leqslant t \leqslant b$, démontrer l'inégalité de Hölder

$$\left| \int_a^b x(t)\, y(t)\, dt \right| \leqslant \sqrt[p]{\int_a^b |x(t)|^p\, dt} \sqrt[q]{\int_a^b |y(t)|^q\, dt}.$$

15. Pour les mêmes fonctions $x(t)$ et $y(t)$, démontrer l'inégalité

$$\sqrt[p]{\int_a^b |x(t)+y(t)|^p\, dt} \leqslant \sqrt[p]{\int_a^b |x(t)|^p\, dt} + \sqrt[p]{\int_a^b |y(t)|^p\, dt}$$

pour $p \geqslant 1$.

16. Soit $p \geqslant 1$, $q \geqslant 1$ et $\frac{1}{p}+\frac{1}{q}=1$. Pour deux vecteurs quelconques $x=\{\xi_1, \ldots, \xi_n\}$ et $y=\{\eta_1, \ldots, \eta_n\}$, démontrer l'inégalité de Hölder

$$\left|\sum_{k=1}^{n} \xi_k \eta_k\right| \leqslant \sqrt[p]{\sum_{k=1}^{n} |\xi_k|^p} \cdot \sqrt[q]{\sum_{k=1}^{n} |\eta_k|^q}.$$

17. Pour les mêmes vecteurs $x=\{\xi_k\}$ et $y=\{\eta_k\}$, démontrer l'inégalité de triangle

$$\sqrt[p]{\sum_{k=1}^{n} |\xi_k+\eta_k|^p} \leqslant \sqrt[p]{\sum_{k=1}^{n} |\xi_k|^p}+\sqrt[p]{\sum_{k=1}^{n} |\eta_k|^p}$$

pour $p \geqslant 1$.

18. Démontrer que l'espace normé l_p (12.33*d*) est complet quel que soit $p \geqslant 1$.

19. On a une suite de fonctions $x_n(t)$ ($n=1, 2, \ldots$) définies et indéfiniment dérivables sur un intervalle $a \leqslant t \leqslant b$. Supposons qu'il existe une suite de constantes A_k ($k=0, 1, 2, \ldots$) telle que

$$|x_n^{(k)}(t)| \leqslant A_k \quad (n=1, 2, \ldots; \quad k=0, 1, 2, \ldots).$$

Extraire une sous-suite $x_{n_m}(t)$ ($m=1, 2, \ldots$) qui, de même que chaque suite des dérivées $x_{n_m}^{(k)}(t)$ de n'importe quel ordre, converge uniformément sur l'intervalle $a \leqslant t \leqslant b$ lorsque $m \to \infty$.

20. Montrer que la quantité $\|x\|_p$ (12.33*c*) ne vérifie pas l'axiome de triangle 12.31*c* si $p<1$.

21. Il existe une variante du théorème d'Arzelà 12.24 qui n'exige pas la continuité des fonctions $x(t)$ ni la compacité (ni même la métrisabilité) de leur ensemble de définition Q. A savoir, soit $\mathrm{P}(Q)$ une famille de fonctions bornées $x(t)$ définies sur un ensemble quelconque Q, à valeurs dans un espace métrique P, cette famille étant métrisée d'après la formule $\rho(x, y)=\sup \rho(x(t), y(t))$. Un ensemble $E \subset \mathrm{P}(Q)$ est précompact si, et seulement si, pour tout $\varepsilon>0$ il existe une partition de l'ensemble Q en un nombre fini de sous-ensembles $Q_1, \ldots, Q_n$ sur chacun desquels la variation de toute fonction de E ne dépasse pas ε.

Historique

Les structures fondamentales de l'analyse sont dégagées à la fin du XIXe et au début du XXe siècle lorsque l'analyse accumule déjà un nombre important de faits et que la nécessité de leur organisation devient impérative. La dépendance linéaire des vecteurs, la dimension (égale à n'importe quel n naturel) figurent déjà chez Grassmann (1846), mais les espaces vectoriels abstraits apparaissent pour la première fois chez Peano (1888).

La théorie des espaces de fonctions continues se développe en Italie durant les années 1870-80 (Volterra, Ascoli, Arzelà, Dini). Le théorème sur les conditions de compacité d'un ensemble de fonctions continues, que l'on appelle d'habitude théorème d'Arzelà, est démontré pour la première fois par Ascoli en 1883. Le théorème de Weierstrass sur l'approximation des fonctions continues par les polynômes date des années 1870; sa généralisation (12.52-12.53) est trouvée par M. Stone en 1936.

L'étape suivante est l'introduction des espaces hilbertiens. Elle s'ouvre avec la construction d'une théorie des équations linéaires intégrales par Volterra (1887) et Fredholm (1900). Hilbert découvre en 1906 une analogie remarquable entre le problème sur les valeurs propres d'opérateurs intégraux et celui de réduction d'une forme quadratique, la résolution du problème pour les opéra-

teurs intégraux s'avérant liée à leur compacité. En 1907, E. Schmidt donne un nouvel exposé de la théorie de Hibert en écrivant les opérateurs intégraux à l'aide de matrices infinies agissant dans l'« espace hilbertien » des suites de carré sommable. Une théorie axiomatique des espaces hilbertiens basée sur la notion de produit scalaire est construite aux environs de 1930 par Stone et von Neumann.

Une autre construction de la théorie des opérateurs compacts qui, en fait, est valable pour n'importe quel espace normé complet (formellement, pour l'espace des fonctions continues) est réalisée par F. Riesz en 1918. La définition abstraite des espaces vectoriels normés apparaît un peu plus tard, en 1920-22, dans les travaux de Banach, Hahn et Wiener. L'école de Banach découvre durant les années 20 les principes fondamentaux de l'analyse fonctionnelle linéaire, y compris le théorème sur l'application ouverte et celui de la borne uniforme (12.74). Les résultats obtenus par cette école et un grand nombre d'applications sont exposés dans [20]. Pourtant, le problème fondamental de la représentation canonique d'un opérateur linéaire quelconque, analogue à la représentation jordanienne d'un opérateur linéaire de dimension finie, attend encore sa solution. Dans ce sens, il y a beaucoup de résultats intéressants et forts concernant les opérateurs dans un espace hilbertien. Hilbert a obtenu dès 1906-1911 un analogue de la représentation diagonale pour les opérateurs symétriques (et hermitiens) non forcément compacts. Le passage à des opérateurs non symétriques est extrêmement lent; les premiers résultats de valeur (liés, pour l'essentiel, aux noms de Livchitz et Keldych) se rapportent à la fin des années 40. Pour l'état actuel de la théorie voir les monographies [4] et [5].

La théorie des algèbres normées dont nous n'avons fourni que les premières notions est construite par Gelfand en 1937-1939; [3] et [8] en donnent un exposé, avec des exemples variés d'applications à l'analyse.

La première approche scientifique de la sommation de suites divergentes est due à Euler (« Institutions du calcul différentiel », 1755). Il ne s'agit encore pas, à cette époque, d'une théorie rigoureuse; de plus, l'utilisation incorrecte de séries divergentes a bien sapé leur prestige. La réforme de Cauchy (1821) chasse pour longtemps les suites et les séries divergentes de l'analyse. La théorie moderne de la sommation de suites est formée au carrefour du XIXe et du XXe siècle (Cesàro, 1880; Voronoï, 1902; Toeplitz, 1911); on peut prendre connaissance de son état actuel dans [21].

Equations différentielles

> Une intelligence qui, pour un instant donné, connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'Analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome; rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux. L'esprit humain offre, dans la perfection qu'il a su donner à l'Astronomie, une faible esquisse de cette intelligence
>
> *Pierre-Simon Laplace*,
> Essai philosophique sur les Probabilités (1795)

§ 13.1. Définitions et exemples

13.11.a. Si une équation par rapport à une fonction inconnue $u = u(t)$, $a \leqslant t \leqslant b$, contient une dérivée (première ou d'un ordre supérieur) de cette fonction, elle est appelée *équation différentielle*. La fonction $u(t)$ peut être cherchée, selon les conditions du problème considéré, soit parmi les fonctions numériques, soit parmi les fonctions vectorielles à valeurs dans un espace n-dimensionnel, soit enfin parmi les fonctions vectorielles à valeurs dans un espace vectoriel normé.

Toute fonction $u(t)$ vérifiant une équation différentielle donnée en est une *solution* ou une *solution particulière*. L'ensemble de toutes les solutions s'appelle *solution générale*.

b. Ainsi, la plus simple équation différentielle

$$u'(t) = 0 \quad (a \leqslant t \leqslant b) \tag{1}$$

a $u(t) = \text{const}$ pour solution générale ; c'est une constante numérique si l'on sait que $u(t)$ prend les valeurs numériques (7.45c), une cons-

tante vectorielle si $u(t)$ prend les valeurs vectorielles, un élément constant d'un espace normé B si $u(t)$ prend les valeurs dans l'espace B (12.61*k*). La solution générale de l'équation différentielle

$$u'(t) = g(t), \tag{2}$$

où $g(t)$ est une fonction donnée (numérique ou vectorielle), peut être mise sous la forme de l'intégrale

$$u(t) = \int_{t_0}^{t} g(\tau)\, d\tau + \text{const},$$

à condition que $g(t)$ soit continue par morceaux (9.32 et 12.63*c*). Les exemples (1) et (2) montrent qu'une équation différentielle ne détermine pas sa solution d'une façon univoque, de sorte que, pour fixer une solution, il faut lui imposer des conditions supplémentaires. D'ordinaire, en tant que condition supplémentaire, on se donne la valeur de la fonction inconnue $u(t)$ pour une certaine valeur $t = t_0 \in [a, b]$. Une fois $u(t_0)$ donné, la solution de l'équation (1) ou (2) est déterminée d'une façon univoque.

c. Une équation plus générale a la forme

$$u'(t) = \Phi(t, u(t)), \tag{3}$$

où $\Phi(t, z)$ est une fonction à valeurs dans le même espace normé B qui contient les valeurs de $u(t)$. Ici on s'interroge sur l'existence d'une solution $u(t)$ et sur son unicité pour $u(t_0)$ donné.

d. Il est utile d'attribuer à (3) un sens « cinématique » dans l'espace B. Figurons-nous un point mobile dans l'espace B dont la position à chaque instant t est $u = u(t)$. Lorsque t varie, le point parcourt dans B une courbe $u = u(t)$ $(a \leqslant t \leqslant b)$. La fonction vectorielle $u(t)$ représente la *loi du mouvement* du point mobile le long de cette courbe. Alors le vecteur $u'(t)$ peut être interprété comme vecteur vitesse du point (limite du rapport du chemin parcouru Δu au temps employé Δt). Le second membre de l'équation (3) fait correspondre à un vecteur $z \in B$, pour tout $t \in [a, b]$ fixe, le vecteur $\Phi(t, z)$. La solution $u(t)$ s'interprète comme loi du mouvement d'un point mobile dans l'espace B lorsque la vitesse du mouvement, à tout instant t et pour toute position $u \in B$, coïncide avec le vecteur $\Phi(t, u)$. Ainsi, à tout instant t, la fonction $\Phi(t, u)$ définit un champ des vecteurs dont chacun détermine la vitesse du mouvement du mobile au point correspondant u de l'espace B. Les solutions de l'équation (3) représentent les trajectoires possibles du point mobile et sont, dans ce cas, appelées *courbes intégrales* de cette équation.

Dans l'espace $R_1 \times B$, on peut attribuer à l'équation (3) un sens purement géométrique. A toute fonction $u = u(t)$, il correspond une courbe dans l'espace $R_1 \times B$ $(t \in R_1,\ u \in B)$. La différentielle

$u'(t)\,dt$ est la partie linéaire principale de l'accroissement de la fonction $u(t)$ lorsque la variable indépendante varie de t à $t+dt$. Par conséquent, dans le cas réel ($B = R_1$), la dérivée $u'(t)$ est interprétée comme coefficient angulaire de la tangente, i.e. comme pente de la courbe $u(t)$, lorsqu'on passe de t à $t+dt$, à des infiniment petits d'ordre supérieur près. Dans le cas général, pour un B quelconque, la dérivée a un sens analogue: la droite $u - u_0 = u'(t_0)(t - t_0)$ est tangente à la courbe $u = u(t)$ pour $t = t_0$, et la quantité $u'(t_0)$ peut être appelée son coefficient angulaire. La fonction $\Phi(t, z)$ détermine à tout point (t_0, z_0) do l'espace $R_1 \times B$ une droite

$$u - z_0 = \Phi(t_0, z_0)(t - t_0), \quad (4)$$

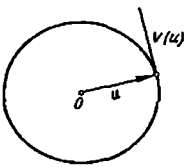


Fig. 13.1.

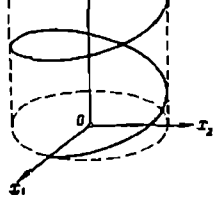


Fig. 13.2.

et l'équation (3) exige que, pour tout $t_0 \in [a, b]$, la courbe $u = u(t)$ ait une tangente se confondant avec la droite (4) déterminée au point $[t_0, u(t_0)]$.

e. En tant qu'exemple, considérons dans l'espace $B = R_2$ l'équation

$$u'(t) = v(u),$$

où $v(u)$ désigne, pour tout $u \in R_2$, le vecteur que l'on obtient en faisant tourner le vecteur u d'un angle droit dans le sens positif.

Du point de vue cinématique, le point mobile doit se déplacer dans le plan R_2 de façon que son vecteur vitesse se confonde avec le vecteur $v(u)$ en tout point u. Il est évident que chaque solution représente un mouvement le long d'une circonférence centrée à l'origine des coordonnées, avec la vitesse égale numériquement au rayon de la circonférence (fig. 13.1).

Du point de vue géométrique, on cherche les courbes dans l'espace tridimensionnel $R_1 \times R_2$ dont la tangente à tout point soit donnée par l'équation

$$u - z_0 = v(z_0)(t - t_0).$$

Les courbes cherchées ont la forme d'hélices contournant l'axe des t (fig. 13.2).

f. Nous verrons plus bas qu'un système d'équations de la forme

$$\left.\begin{aligned} u_1'(t) &= \Phi_1(t, u_1(t), \ldots, u_n(t)), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ u_n'(t) &= \Phi_n(t, u_1(t), \ldots, u_n(t)), \end{aligned}\right\} \tag{5}$$

ainsi qu'une équation du n-ième ordre

$$u^{(n)} = \Phi(t, u(t), u'(t), \ldots, u^{(n-1)}(t)) \tag{6}$$

se réduisent à une équation du type (3).

g. Les problèmes d'existence de solutions des équations différentielles et de leur unicité dans certaines conditions supplémentaires nous intéresseront le long de tout le chapitre ; pour le moment, nous considérons quelques cas très simples où la solution s'obtient sous une forme explicite.

13.12. Soit une équation de la forme

$$u'(t) = A(t)\, u(t) \quad (a \leqslant t \leqslant b). \tag{1}$$

Une telle équation s'appelle *équation linéaire homogène*. Supposons d'abord que la fonction cherchée $u(t)$ soit une fonction numérique et le coefficient $A(t)$ une fonction numérique continue donnée. La valeur $u_0 = u(t_0)$ est également supposée donnée. La fonction $u(t) \equiv 0$ est une solution évidente de l'équation (1), mais elle ne satisfait pas, pour $u_0 \neq 0$, à la condition initiale. Cherchons d'autres solutions. Si $u(t)$ est une solution non identiquement nulle, alors il existe un intervalle dans lequel $u(t) \neq 0$, par exemple $u(t) > 0$. En divisant (1) par $u(t)$, on obtient

$$\frac{u'(t)}{u(t)} \equiv [\ln u(t)]' = A(t),$$

d'où

$$\ln u(t) = \int_{t_0}^{t} A(\tau)\, d\tau + C.$$

En posant $t = t_0$ on a $C = \ln u_0$; définitivement, en se débarrassant des logarithmes on trouve

$$u(t) = e^{\int_{t_0}^{t} A(\tau)\, d\tau} u_0. \tag{2}$$

La vérification directe montre que (2) est réellement une solution de l'équation (1) qui ne dépend plus du signe de $u(t)$. Remarquons que la solution (2) est définie pour tout $t \in [a, b]$ (et non nulle partout si $u_0 \neq 0$). Donc, *pour la condition initiale* $u(t_0) = u_0$, *l'équation* (1) *possède la solution* (2). Si $A(t) \equiv A$ est une constante, alors

la solution (2) prend une forme particulièrement simple

$$u(t) = e^{(t-t_0)A} u_0. \tag{3}$$

13.13. Revenons à l'équation

$$u'(t) = A(t)\, u(t) \quad (a \leqslant t \leqslant b) \tag{1}$$

en supposant cette fois que $u(t)$ soit une fonction vectorielle à valeurs dans un espace de Banach B ; le coefficient $A(t)$ est supposé un opérateur linéaire continu qui applique, pour tout $t \in [a, b]$, l'espace B dans lui-même et dépend continûment du paramètre t. Ici les raisonnements 13.12 ne sont plus valables car la division par $u(t)$ n'a plus de sens. Il s'avère cependant que l'on peut attribuer un sens rigoureux aux résultats définitifs 13.12 (2) ou (3).

Supposons d'abord que l'opérateur $A(t) \equiv A$ ne dépende pas de t ; le cas général sera étudié dans 13.19.

Nous considérons $e^{(t-t_0)A}$ comme fonction de l'opérateur A au sens de 12.86*c* :

$$e^{(t-t_0)A} = \sum_{n=0}^{\infty} \frac{(t-t_0)^n A^n}{n!}. \tag{2}$$

Cette fonction est définie pour tous les t réels et prend ses valeurs dans l'espace L (B) des opérateurs linéaires bornés dans B. La série (2) peut être dérivée terme à terme par rapport à t (12.66), ce qui fournit

$$\frac{d}{dt}\, e^{(t-t_0)A} = \sum_{n=0}^{\infty} \frac{n(t-t_0)^{n-1} A^n}{n!} = A e^{(t-t_0)A}.$$

On en conclut que $u(t) = e^{(t-t_0)A} u(t_0)$ est réellement une solution de l'équation (1). Pour $t = t_0$, cette solution donne, évidemment, le vecteur $u = u(t_0)$. Ainsi, pour un $u(t_0)$ donné, on a une solution de l'équation homogène (1), qui est de la forme

$$u(t) = e^{(t-t_0)A} u(t_0). \tag{3}$$

Pour démontrer l'unicité de la solution obtenue, démontrons d'abord le lemme suivant :

Lemme. *Si $B(t)$ est une fonction opératorielle fortement dérivable* (i.e., pour tout $x \in X$, on a la relation limite $B'(t)\, x = \lim_{\Delta t \to 0} \frac{B(t+\Delta t) - B(t)}{\Delta t}\, x$) *et si $x(t)$ est une fonction vectorielle dérivable, alors la fonction vectorielle $y(t) = B(t)\, x(t)$ est elle aussi dérivable et l'on a*

$$y'(t) = B(t)\, x'(t) + B'(t)\, x(t).$$

Démonstration. On a

$$\frac{B(t+\Delta t)\,x(t+\Delta t)-B(t)\,x(t)}{\Delta t}=$$

$$=B(t+\Delta t)\,\frac{x(t+\Delta t)-x(t)}{\Delta t}+\frac{B(t+\Delta t)-B(t)}{\Delta t}\,x(t).$$

Le premier terme à droite tend vers la limite $B(t)\,x'(t)$ lorsque $\Delta t \to 0$ (12.74 *e-f*), le deuxième tend vers $B'(t)\,x(t)$ par hypothèse, d'où le lemme.

A présent, démontrons que (3) *est la solution unique de l'équation* (1) *avec la valeur* $u(t_0)$ *donnée*. Soit $u(t)$ une solution quelconque de l'équation (1) avec $u(t_0)$ donné. Introduisons une nouvelle fonction inconnue $v(t)$ d'après la formule $u(t)=e^{(t-t_0)A}v(t)$ ou, ce qui revient au même, $v(t)=e^{-(t-t_0)A}u(t)$. En portant $u(t)$ dans l'équation (1) et en utilisant le lemme on trouve

$$u'(t)=Ae^{(t-t_0)A}v(t)+e^{(t-t_0)A}v'(t)=Ae^{(t-t_0)A}v(t),$$

d'où

$$e^{(t-t_0)A}v'(t)=0\ ;$$

en multipliant par $e^{-(t-t_0)A}$ on obtient $v'(t)=0$. Il en résulte que

$$v(t)\equiv v(t_0)=u(t_0),$$

donc la solution $u(t)$ a la forme (3), ce qu'il fallait démontrer.

13.14. Comment sera la solution 13.13 (3) en cas d'espace n-dimensionnel réel R_n? Pour simplifier, posons $t_0=0$. Choisissons dans l'espace R_n une base $e_1, \ldots, e_n$. Développons la fonction vectorielle $u(t)$ suivant cette base; soit

$$u(t)=\sum_{k=1}^{n}u_k(t)\,e_k.$$

Alors

$$u'(t)=\sum_{k=1}^{n}u_k'(t)\,e_k,$$

et l'équation vectorielle 13.12 (1) se met sous la forme du système d'équations scalaires:

$$\left.\begin{aligned}\frac{du_1(t)}{dt}&=a_{11}u_1(t)+\ldots+a_{1n}u_n(t),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ \frac{du_n(t)}{dt}&=a_{n1}u_1(t)+\ldots+a_{nn}u_n(t)\end{aligned}\right\}\qquad(1)$$

avec une matrice réelle constante $A=\|a_{jk}\|$. La solution est le résultat de l'application de l'opérateur e^{tA} au vecteur initial $u_0=u(0)$.

Il s'avère que la solution cherchée peut être mise sous une forme explicite suffisamment simple si l'on choisit pour vecteurs initiaux u_0 les *vecteurs d'une base jordanienne de la matrice A* (12.17*f*). Pour de tels vecteurs, nous introduisons les désignations suivantes.

a) Le vecteur de base associé à une case jordanienne à un seul élément λ_j sera désigné par f_j.

b) Les vecteurs de base associés à une $m \times m$-case jordanienne

$$\begin{Vmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_j & 1 \\ & & & & \lambda_j \end{Vmatrix} \tag{2}$$

(λ_j étant réel) seront désignés par $f_j^1, \ldots, f_j^m$. Les vecteurs de base associés à une 2×2-case

$$\begin{Vmatrix} \sigma_j & -\tau_j \\ \tau_j & \sigma_j \end{Vmatrix}, \qquad \lambda_j = \sigma_j + i\tau_j \tag{3}$$

seront désignés par h_j, g_j; enfin, les vecteurs de base associés à une $2m \times 2m$-case

$$\begin{Vmatrix} \sigma_j & -\tau_j & 1 & 0 & \ldots \\ \tau_j & \sigma_j & 0 & 1 & \ldots \\ & & \sigma_j & -\tau_j & \ldots \\ & & \tau_j & \sigma_j & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{Vmatrix} \tag{4}$$

seront désignés par $h_j^1, g_j^1, \ldots, h_j^m, g_j^m$. Rappelons que, dans tous les cas considérés, les nombres λ_j sont des racines de l'équation caractéristique

$$\begin{Vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{Vmatrix} = 0,$$

les nombres σ_j et τ_j étant respectivement les parties réelle et imaginaire des racines complexes de cette équation (12.17*f*).

Toute case de la matrice jordanienne définit un sous-espace invariant de l'opérateur A (de dimension 1, m, 2, $2m$ respectivement). L'opérateur e^{tA} appliqué à un vecteur de ce sous-espace donne un autre vecteur de ce sous-espace. Les solutions répondant aux vecteurs initiaux $f_j, f_j^s, h_j, g_j, h_j^s, g_j^s$ ($s = 1, \ldots, m$) seront désignées respectivement par $f_j(t)$, $f_j^s(t)$, $h_j(t)$, $g_j(t)$, $h_j^s(t)$, $g_j^s(t)$. Dans l'espace invariant unidimensionnel engendré par le vecteur f_j, l'opérateur A se réduit à la multiplication par λ_j et l'opérateur e^{tA} à la multiplication par $e^{t\lambda_j}$. On en tire

$$f_j(t) = e^{t\lambda_j} f_j. \tag{5}$$

Pour l'espace invariant m-dimensionnel correspondant à la case jordanienne (2), nous avons (12.19e)

$$e^{tA} = \left\| \begin{matrix} e^{\lambda_j t} & te^{\lambda_j t} & \frac{t^2}{2} e^{\lambda_j t} & \dots & \frac{t^{m-1}}{(m-1)!} e^{\lambda_j t} \\ 0 & e^{\lambda_j t} & te^{\lambda_j t} & \dots & \frac{t^{m-2}}{(m-2)!} e^{\lambda_j t} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & e^{\lambda_j t} \end{matrix} \right\|.$$

Par conséquent,

$$\left.\begin{aligned} f_j^1(t) &= e^{tA} f_j^1 = e^{\lambda_j t} f_j^1, \\ f_j^2(t) &= e^{tA} f_j^2 = e^{\lambda_j t}(t f_j^1 + f_j^2), \\ &\dots \\ f_j^m(t) &= e^{tA} f_j^m = e^{\lambda_j t}\left(\frac{t^{m-1}}{(m-1)!} f_j^1 + \frac{t^{m-2}}{(m-2)!} f_j^2 + \dots + f_j^m\right). \end{aligned}\right\} \quad (6)$$

c) Pour la 2×2-case jordanienne (3), nous avons (12.19h)

$$e^{tA} = e^{\sigma t} \left\| \begin{matrix} \cos t\tau & -\sin t\tau \\ \sin t\tau & \cos t\tau \end{matrix} \right\|,$$

par conséquent,

$$\left.\begin{aligned} h_j(t) &= e^{\sigma t}[\cos t\tau \cdot h_j + \sin t\tau \cdot g_j], \\ g_j(t) &= e^{\sigma t}[-\sin t\tau \cdot h_j + \cos t\tau \cdot g_j]. \end{aligned}\right\} \quad (7)$$

d) Pour la $2m \times 2m$-case jordanienne (4), nous avons (12.19h)

$$e^{tA} = e^{\sigma t} \left\| \begin{matrix} \cos t\tau & -\sin t\tau & t\cos t\tau & -t\sin t\tau & \dots \\ \sin t\tau & \cos t\tau & t\sin t\tau & t\cos t\tau & \dots \\ & & \cos t\tau & -\sin t\tau & \dots \\ & & \sin t\tau & \cos t\tau & \dots \\ \dots & \dots & \dots & \dots & \dots \\ & & & & \dots \cos t\tau \quad -\sin t\tau \\ & & & & \dots \sin t\tau \quad \cos t\tau \end{matrix} \right\|$$

de sorte que

$$\left.\begin{aligned} h_j^1(t) &= e^{\sigma t}[\cos \tau t \cdot h_j^1 + \sin \tau t \cdot g_j^1], \\ g_j^1(t) &= e^{\sigma t}[-\sin \tau t \cdot h_j^1 + \cos \tau t \cdot g_j^1], \\ &\dots \\ h_j^m(t) &= e^{\sigma t}\left[\frac{t^{m-1}}{(m-1)!}(\cos \tau t \cdot h_j^1 + \sin \tau t \cdot g_j^1) + \dots \right. \\ &\qquad \left. \dots + \cos \tau t \cdot h_j^m + \sin \tau t \cdot g_j^m\right], \\ g_j^m(t) &= e^{\sigma t}\left[\frac{t^{m-1}}{(m-1)!}(-\sin \tau t \cdot h_j^1 + \cos \tau t \cdot g_j^1) + \dots \right. \\ &\qquad \left. \dots + (-\sin \tau t \cdot h_j^m + \cos \tau t \cdot g_j^m)\right]. \end{aligned}\right\} \quad (8)$$

13.15. Nous avons construit n solutions particulières différentes de l'équation 13.12(1) dans l'espace R_n qui correspondent à n vecteurs d'une base jordanienne de la matrice A pris en tant que vecteurs initiaux. Dans la base de départ $e_1, \ldots, e_n$ par rapport à laquelle l'équation 13.12(1) a la forme du système 13.14(1), chacune de ces n solutions est représentée à l'aide de n fonctions scalaires (coordonnées). Soient, par exemple,

$$f_j(t)=\sum_{k=1}^{n} u_{jk}(t)\,e_k, \quad h_j(t)=\sum_{k=1}^{n} v_{jk}(t)\,e_k, \quad g_j(t)=\sum_{k=1}^{n} w_{jk}(t)\,e_k,$$

$$f_j^s(t)=\sum_{k=1}^{n} u_{jk}^s(t)\,e_k, \quad h_j^s(t)=\sum_{k=1}^{n} v_{jk}^s(t)\,e_k, \quad g_j^s(t)=\sum_{k=1}^{n} w_{jk}^s(t)\,e_k,$$

puis

$$f_j=\sum u_{jk}e_k, \quad h_j=\sum v_{jk}e_k, \quad g_j=\sum w_{jk}e_k,$$

$$f_j^s=\sum u_{jk}^s e_k, \quad h_j^s=\sum v_{jk}^s e_k, \quad g_j^s=\sum w_{jk}^s e_k.$$

Nous avons

$$f_j(t)=e^{\lambda_j t}f_j=e^{\lambda_j t}\sum u_{jk}e_k=\sum u_{jk}(t)\,e_k,$$

d'où

$$u_{jk}(t)=e^{\lambda_j t}u_{jk}. \tag{1}$$

D'une façon analogue, nous obtenons

$$u_{jk}^s(t)=e^{\lambda_j t}\left[\frac{t^{s-1}}{(s-1)!}u_{jk}^1+\ldots+u_{jk}^s\right], \tag{2}$$

$$\left.\begin{aligned} v_{jk}(t)&=e^{\sigma_j t}[\cos\tau_j t\cdot v_{jk}+\sin\tau_j t\cdot w_{jk}],\\ w_{jk}(t)&=e^{\sigma_j t}[-\sin\tau_j t\cdot v_{jk}+\cos\tau_j t\cdot w_{jk}], \end{aligned}\right\} \tag{3}$$

$$\left.\begin{aligned} v_{jk}^s(t)&=e^{\sigma_j t}\left[\frac{t^{s-1}}{(s-1)!}(\cos\tau_j t\cdot v_{jk}^1+\sin\tau_j t\cdot w_{jk}^1)+\ldots\right.\\ &\qquad\left.\ldots+(\cos\tau_j t\cdot v_{jk}^s+\sin\tau_j t\cdot w_{jk}^s)\right],\\ w_{jk}^s(t)&=e^{\sigma_j t}\left[\frac{t^{s-1}}{(s-1)!}(-\sin\tau_j t\cdot v_{jk}^1+\cos\tau_j t\cdot w_{jk}^1)+\ldots\right.\\ &\qquad\left.\ldots+(-\sin\tau_j t\cdot v_{jk}^s+\cos\tau_j t\cdot w_{jk}^s)\right]. \end{aligned}\right\} \tag{4}$$

13.16. Les formules 13.14(5)-(8) peuvent être appliquées à l'étude des courbes intégrales $u=u(t)$ et de leur comportement asymptotique pour $t\to\infty$. Il est commode d'utiliser l'interprétation cinématique 13.11*d*.

a) Pour un vecteur initial du type f_j (13.14a), la solution $f_j(t)$ est ou bien le vecteur constant f_j lorsque $\lambda_j=0$, ou bien un vecteur s'éloignant du zéro selon la loi exponentielle (pour $t\to\infty$) le long de l'axe f_j lorsque $\lambda_j>0$, ou enfin un vecteur s'en approchant selon la même loi le long du même axe lorsque $\lambda_j<0$.

b) Supposons que le vecteur initial u_0 soit un vecteur du type f_j^k, i.e. l'un des vecteurs de base d'un sous-espace invariant m-dimensionnel associé à une case jordanienne 13.14(2). Si l'on utilise la formule correspondante du groupe

13.14(6), on obtient la solution

$$f_j^k(t) = e^{t\lambda_j}\left[\frac{t^{k-1}}{(k-1)!}f_j^1 + \frac{t^{k-2}}{(k-2)!}f_j^2 + \ldots + \frac{t}{1!}f_j^{k-1} + f_j^k\right].$$

Nous voyons que le rayon vecteur de la courbe intégrale correspondante, qui se confond à l'instant initial avec le vecteur f_j^k, acquiert avec le temps les composantes non nulles suivant les vecteurs $f_j^{k-1}, \ldots, f_j^1$ de sorte que, pour t grand, la composante suivant le vecteur f_j^1 devient dominante. Si $\lambda_j \geqslant 0$, $k > 1$, alors la courbe s'éloigne, pour $t \to \infty$, de l'origine des coordonnées, et sa tangente (on peut le voir en dérivant) devient à la limite parallèle au vecteur f_j^1. Pour $\lambda_j = 0$, la vitesse du point mobile s'éloignant de l'origine des coordonnées varie selon la loi de puissance, pour $\lambda_j > 0$, selon la loi exponentielle. Si $\lambda_j < 0$, alors, pour $t \to \infty$, la courbe s'approche de l'origine des coordonnées; la composante suivant f_j^1 étant dominante, la courbe entre, lorsque $t \to \infty$, dans un cône aussi étroit que l'on veut, de sommet à l'origine des coordonnées et d'axe dirigé le long de f_j^1; ceci veut dire que la position limite de sa tangente se confond avec la tangente au vecteur f_j^1.

c) Supposons que le vecteur initial $u(t_0)$ soit un vecteur h_j ou bien g_j dans un sous-espace invariant bidimensionnel H_2 correspondant à une 2×2-case jordanienne que l'on a considérée dans 13.14c). Alors les formules 13.14(7) montrent que la solution $u(t)$ décrit dans le plan H_2:

une ellipse centrée à l'origine des coordonnées si $\sigma_j = 0$;

une spirale s'éloignant de l'origine si $\sigma_j > 0$;

une spirale qui se rapproche de l'origine des coordonnées et tend vers l'origine lorsque $t \to \infty$ si $\sigma_j < 0$.

d) Supposons que le vecteur initial $u(t_0)$ soit un vecteur h_j^k ou g_j^k dans un sous-espace invariant $2m$-dimensionnel H_{2m} associé à une case jordanienne de dimensions $2m \times 2m$ indiquée dans 13.14d). Alors la solution $u(t)$ décrit dans l'espace H_{2m} l'une des courbes suivantes:

si $\sigma_j \geqslant 0$, une spirale s'éloignant de l'origine dont la tangente tend à devenir parallèle, pour $t \to \infty$, au plan du premier couple des vecteurs de base h_j^1, g_j^1;

si $\sigma_j < 0$, une spirale qui se rapproche de l'origine des coordonnées et y tend, pour $t \to \infty$, en devenant tangente au plan du premier couple des vecteurs de base.

e) Dans le cas général où le vecteur $u(t_0)$ possède plusieurs composantes suivant des vecteurs d'une base jordanienne, le mouvement correspondant est la somme géométrique des mouvements considérés.

13.17. Une équation scalaire linéaire du n-ième ordre

$$y^{(n)}(t) = a_1(t)\,y(t) + \ldots + a_n(t)\,y^{(n-1)}(t) \tag{1}$$

peut être mise sous la forme d'un système du premier ordre en posant

$$y(t) = u_1(t), \quad y'(t) = u_2(t), \quad \ldots, \quad y^{(n-1)}(t) = u_n(t). \tag{2}$$

Avec cette substitution nous avons

$$\left.\begin{aligned} u_1'(t) &= u_2(t), \\ u_2'(t) &= u_3(t), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ u_n'(t) &= a_1(t)\,u_1(t) + a_2(t)\,u_2(t) + \ldots + a_n(t)\,u_n(t). \end{aligned}\right\} \tag{3}$$

Inversement, toute solution $(u_1(t), \ldots, u_n(t))$ du système (3) permet de déterminer d'après les formules (2) une fonction $y(t) = u_1(t)$ et ses dérivées; la dernière équation (3) montre que la fonction $y(t)$ vérifie l'équation (1).

En supposant $a_1(t) = a_1, \ldots, a_n(t) = a_n$ constants, appliquons les résultats 13.14. La matrice de l'opérateur A a une forme particulière

$$A = \begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ a_1 & a_2 & a_3 & \ldots & a_{n-1} & a_n \end{Vmatrix}.$$

Soit $f = \sum_{j=1}^{n} \xi_j e_j$ un vecteur propre de l'opérateur A qui correspond à une valeur propre λ. Nous avons $Af = \lambda f$, ou bien, en coordonnées $\xi_1, \ldots, \xi_n$,

$$\xi_2 = \lambda\xi_1,$$
$$\xi_3 = \lambda\xi_2,$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$a_1\xi_1 + a_2\xi_2 + a_3\xi_3 + \ldots + a_n\xi_n = \lambda\xi_n.$$

En posant $\xi_1 = 1$ trouvons successivement

$$\xi_1 = 1,\ \xi_2 = \lambda,\ \xi_3 = \lambda^2,\ \ldots,\ \xi_n = \lambda^{n-1},$$
$$a_1 + a_2\lambda + \ldots + a_n\lambda^{n-1} = \lambda^n. \tag{4}$$

C'est l'équation *caractéristique* de l'équation (1). Elle s'en déduit en remplaçant $y^{(k)}$ par λ^k $(k = 0, 1, \ldots, n)$.

Ainsi, les racines caractéristiques de la matrice A sont racines de l'équation caractéristique (4). Un vecteur propre associé à une valeur propre λ est colinéaire au vecteur $(1, \lambda, \lambda^2, \ldots, \lambda^{n-1})$, donc défini d'une façon univoque (à la colinéarité près). Par conséquent, dans le cas où λ est une racine de multiplicité m, on voit apparaître une case jordanienne réelle ou complexe à m lignes et m colonnes. (En d'autres termes, les exposants des puissances des diviseurs élémentaires sont dans ce cas égaux aux multiplicités des racines, et le polynôme minimal de la matrice A se confond avec son polynôme caractéristique [14; ch. 6].)

Conformément à 13.14, on peut écrire n solutions particulières différentes du système (3) qui correspondent à n vecteurs d'une base jordanienne pris pour vecteurs initiaux. Bornons-nous à expliciter les premières composantes de ces solutions vu que, en vertu des formules (2), il nous faut justement la première composante $u_1(t) = y(t)$:

$$u_{j1}(t) = u_{j1}e^{\lambda_j t}, \tag{5}$$

pour toute racine réelle simple λ_j;

$$u_j^s(t) = e^{\lambda_j t}\left[\frac{t^{s-1}}{(s-1)!}u_j^1 + \ldots + u_j^s\right], \quad s = 1, \ldots, m, \tag{6}$$

pour toute racine réelle λ_j de multiplicité m;

$$\left.\begin{aligned} v_j(t) &= e^{\sigma_j t}[v_{j1}\cos\tau_j t + w_{j1}\sin\tau_j t], \\ w_j(t) &= e^{\sigma_j t}[-v_{j1}\sin\tau_j t + w_{j1}\cos\tau_j t], \end{aligned}\right\} \tag{7}$$

pour toute racine simple complexe $\lambda_j = \sigma_j + i\tau_j$;

$$\left.\begin{aligned} v_j^s(t) &= e^{\sigma_j t}\left[\frac{t^{s_j-1}}{(s_j-1)!}(v_j^1\cos\tau_j t + w_j^1\sin\tau_j t) + \ldots\right. \\ &\left.\ldots + (v_j^s\cos\tau_j t + w_j^s\sin\tau_j t)\right], \quad s_j = 1, \ldots, m_j, \\ w_j^s(t) &= e^{\sigma_j t}\left[\frac{t^{s_j-1}}{(s_j-1)!}(-v_j^1\sin\tau_j t + w_j^1\cos\tau_j t) + \ldots\right. \\ &\left.\ldots + (-v_j^s\sin\tau_j t + w_j^s\cos\tau_j t)\right], \quad s_j = 1, \ldots, m_j, \end{aligned}\right\} \tag{8}$$

pour toute racine complexe $\lambda_j = \sigma_j + i\tau_j$ de multiplicité m_j.

13.18. En remplaçant les solutions obtenues par certaines de leurs combinaisons linéaires, nous sommes en mesure de signaler les n solutions suivantes :

a) Pour toute racine réelle simple λ_j, nous avons la solution $e^{\lambda_j t}$.

b) Pour toute racine réelle m-uple λ_j, nous avons m solutions :

$$e^{\lambda_j t}, \quad te^{\lambda_j t}, \quad \ldots, \quad t^{m-1}e^{\lambda_j t}.$$

c) Pour tout couple de racines complexes simples $\lambda_j = \sigma_j \pm i\tau_j$, $\tau_j \neq 0$, nous avons deux solutions :

$$e^{\sigma_j t}\cos\tau_j t, \quad e^{\sigma_j t}\sin\tau_j t.$$

d) Pour tout couple de racines complexes m-uples $\lambda_j = \sigma_j \pm i\tau_j$, $\tau_j \neq 0$, nous avons $2m$ solutions :

$$e^{\sigma_j t}\cos\tau_j t,\ e^{\sigma_j t}\sin\tau_j t,\ te^{\sigma_j t}\cos\tau_j t,\ te^{\sigma_j t}\sin\tau_j t,\ \ldots$$

$$\ldots,\ t^{m-1}e^{\sigma_j t}\cos\tau_j t,\ t^{m-1}e^{\sigma_j t}\sin\tau_j t.$$

13.19. Envisageons à présent le cas général où l'opérateur $A(t)$ dans l'équation

$$u'(t) = A(t)\,u(t) \quad (a \leqslant t \leqslant b) \tag{1}$$

dépend réellement du paramètre t ; dans le cas unidimensionnel, nous avons alors la formule 13.12 (2) :

$$u(t) = e^{\int_{t_0}^{t} A(\tau)\,d\tau}\,u_0. \tag{2}$$

Nous pouvons, bien entendu, former l'opérateur $W(t) = \int_{t_0}^{t} A(\tau)\,d\tau$, puis l'expression

$$e^{\int_{t_0}^{t} A(\tau)\,d\tau}\,u(t_0) = e^{W(t)}\,u(t_0),$$

mais, en général, elle n'est plus une solution de l'équation (1). En effet, si l'on essaie de dériver l'expression $e^{W(t)}$ par rapport à t,

on se heurte à la difficulté suivante : il n'est plus possible de se servir de l'égalité

$$e^{W(t)+h\widetilde{A}(t;\,t+h)} = e^{W(t)}\,e^{h\widetilde{A}(t;\,t+h)}$$

pour transformer la différence

$$e^{W(t+h)} - e^{W(t)} = e^{W(t)+h\widetilde{A}(t;\,t+h)} - e^{W(t)},$$

où $\widetilde{A}(t;\,t+h)$ désigne la valeur moyenne de l'opérateur $A(\tau)$ pour $\tau \in (t;\,t+h)$, parce que la relation $e^{A+B} = e^A e^B$, valable pour les opérateurs A et B qui commutent, n'a, en général, plus lieu pour les opérateurs qui ne commutent pas. Dans le cas général, les opérateurs $W(t)$ et $\widetilde{A}(t;\,t+h)$ ne commutent pas. Donc, la dérivée de $e^{W(t)}$ n'est en général pas égale à $e^{W(t)}W'(t)$.

Tout de même, à un certain sens, l'expression (2) peut être considérée comme solution de l'équation (1).

Introduisons la notion d'*intégrale multiplicative*.

Soit $\Pi = \{a = t_0 \leqslant \xi_0 \leqslant t_1 \leqslant \xi_1 \leqslant \ldots \leqslant t_n = t\}$ une partition de l'intervalle $a \leqslant \tau \leqslant t$, à points marqués $\xi_0, \ldots, \xi_{n-1}$. Formons l'opérateur

$$e^{A(\xi_{n-1})\Delta t_{n-1}} e^{A(\xi_{n-2})\Delta t_{n-2}} \ldots e^{A(\xi_0)\Delta t_0}. \tag{3}$$

Si les opérateurs $A(\xi)$, pour des ξ différents, commutaient, on pourrait mettre cet opérateur sous la forme

$$e^{A(\xi_0)\Delta t_0 + \ldots + A(\xi_{n-1})\Delta t_{n-1}}.$$

Pour $d(\Pi) = \max \Delta t_k \to 0$, la dernière expression tend vers la limite $e^{\int_{t_0}^{t} A(\tau)\,d\tau}$. Or, nous avons vu que cet opérateur, pour $A(\xi)$ non commutants, n'est pas une solution de l'équation (1). Il s'avère que, pour $A(\xi)$ non commutants, une solution peut être représentée par l'opérateur qui s'obtient de (3) lorsque $d(\Pi) \to 0$ (cf. exercice 16). Cet opérateur limite est désigné par

$$e^{\int_{t_0}^{t} A(\tau)\,d\tau} \tag{4}$$

et appelé *intégrale multiplicative*.

À des infiniment petits d'ordre supérieur près, on peut écrire

$$e^{A(\xi_j)\Delta t_j} \approx I + A(\xi_j)\,\Delta t_j.$$

On peut démontrer (cf. exercice 14) que la même limite (4) s'obtient si l'on part des produits

$$[I + A(\xi_{n-1})\,\Delta\xi_{n-1}]\,[I + A(\xi_{n-2})\,\Delta\xi_{n-2}] \ldots [I + A(\xi_0)\,\Delta\xi_0].$$

Pour cette raison, l'intégrale multiplicative est parfois désignée par

$$\prod_{t_0}^{t} [I + A(t)]\,dt.$$

Les intégrales multiplicatives peuvent être utilisées pour des estimations de solutions. Cependant, par rapport au théorème général d'existence, leur champ d'applications n'est pas important; c'est pourquoi nous ne donnons pas ici les démonstrations des propositions énoncées.

§ 13.2. Théorème du point fixe

Dans la suite de ce chapitre nous considérerons les théorèmes fondamentaux sur l'existence et l'unicité de solutions des équations différentielles ordinaires. Tous ces théorèmes sont basés sur un principe géométrique important de l'analyse qui s'appelle *principe du point fixe*.

13.21. Soient M un ensemble et A une application de cet ensemble dans lui-même, i.e. une loi selon laquelle on fait correspondre à tout point $x \in M$ un point $y = A(x)$ de M.

Définition. Tout point $x \in M$ que l'application A transforme en lui-même, de sorte que $A(x) = x$, s'appelle *point fixe de l'application* A.

Ainsi, lorsqu'il s'agit de l'application d'un cercle plan M dans lui-même moyennant une rotation de 90° autour du centre, le seul point fixe est le centre du cercle. Si l'application du cercle est l'homothétie de centre O et de rapport 1:2 suivie d'une translation jusqu'au contact avec la circonférence initiale (fig. 13.3), alors le point de contact P en est un point fixe (bien qu'il ne le soit pas pour l'homothétie ni pour la translation en question; ce n'est pas le chemin mais le résultat qui importe). Et lorsqu'il s'agit de l'application d'une circonférence dans elle-même par une rotation de 90°, elle n'a pas de point fixe.

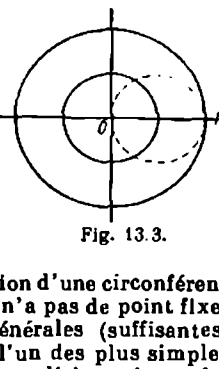


Fig. 13.3.

On a intérêt à établir des conditions générales (suffisantes) d'existence de points fixes. Nous exposons ici l'un des plus simples théorèmes qui garantissent, dans certaines conditions imposées à l'ensemble M et à l'application A, l'existence et l'unicité du point fixe.

13.22. Supposons que M soit un espace métrique.

Définition. Une application A de l'espace métrique M dans lui-même est dite *contractante* s'il existe une constante θ, $0 \leqslant \theta < 1$, telle que, pour deux points quelconques y, z de l'espace M, on a l'inégalité

$$\rho(A(y), A(z)) \leqslant \theta\rho(y, z).$$

Théorème (principe du point fixe de Picard-Banach). *Toute application contractante A d'un espace métrique complet* M *dans lui-même admet un point fixe et un seul.*

Démonstration. A partir d'un point quelconque $x_0 \in M$, construisons la suite des points

$$x_1 = A(x_0),\ x_2 = A(x_1) = A^2(x_0),\ \ldots,\ x_n = A(x_{n-1}) = A^n(x_0),\ \ldots$$

Cette suite est de Cauchy dans M.

En effet, pour tout $n \geqslant 1$, on a

$$\rho(x_n, x_{n+1}) = \rho(A^n(x_0), A^{n+1}(x_0)) \leqslant$$
$$\leqslant \theta\rho(A^{n-1}(x_0), A^n(x_0)) \leqslant \theta^n \rho(x_0, x_1),$$

donc

$$\rho(x_n, x_{n+p}) \leqslant \rho(x_n, x_{n+1}) + \rho(x_{n+1}, x_{n+2}) + \ldots + \rho(x_{n+p-1}, x_{n+p}) \leqslant$$
$$\leqslant [\theta^n + \theta^{n+1} + \ldots + \theta^{n+p-1}]\,\rho(x_0, x_1) \leqslant$$
$$\leqslant \theta^n (1 + \theta + \theta^2 + \ldots)\,\rho(x_0, x_1) = \frac{\theta^n}{1-\theta}\rho(x_0, x_1); \qquad (1)$$

cette quantité devient arbitrairement petite pour n suffisamment grand. M étant complet, il existe une limite

$$x = \lim_{n\to\infty} x_n \in M.$$

Montrons que x est un point fixe. Nous avons, pour $n \geqslant 1$,

$$\rho(A(x), x_n) = \rho(A(x), A(x_{n-1})) \leqslant \theta\rho(x, x_{n-1}) \to 0,$$

d'où

$$A(x) = \lim_{n\to\infty} x_n = x,$$

de sorte que x est bien un point fixe.

Supposons qu'il existe un autre point fixe y, de sorte que l'on ait $A(x) = x$ et $A(y) = y$. Alors

$$\rho(x, y) = \rho(A(x), A(y)) \leqslant \theta\rho(x, y).$$

Si $\rho(x, y) \neq 0$, on peut diviser l'égalité par $\rho(x, y)$ en arrivant à la contradiction $1 \leqslant \theta < 1$. Par conséquent, $\rho(x, y) = 0$, $x = y$ et il n'existe aucun point fixe autre que x. Le théorème est démontré.

13.23. Points fixes de deux applications contractantes. Deux applications $A(y)$ et $B(y)$ d'un espace métrique M dans lui-même sont dites *e-proches* si, pour tout $y \in M$, on a

$$\rho(A(y), B(y)) \leqslant e.$$

Lemme. *Soient $A(y)$ et $B(y)$ deux applications contractantes dans un espace métrique complet* M, *telles que*

$$\rho(A(y), A(z)) \leqslant \theta_A \rho(y, z), \quad \rho(B(y), B(z)) \leqslant \theta_B \rho(y, z),$$

où $\theta_A < 1$, $\theta_B < 1$, *et soit* $\theta = \max(\theta_A, \theta_B)$. *Si les applications* A *et* B *sont* ε*-proches, leurs points fixes sont distants l'un de l'autre d'au plus* $\varepsilon/(1-\theta)$.

Démonstration. Soit y_0 un point fixe de l'application A. Le point fixe z_0 de l'application B peut être obtenu, d'après la construction 13.22, comme limite de la suite $y_0, B(y_0), B^2(y_0), \ldots$ En vertu de l'inégalité 13.22 (1)

$$\rho(y_0, B^n(y_0)) \leqslant \frac{1}{1-\theta} \rho(y_0, B(y_0)) = \frac{1}{1-\theta} \rho(A(y_0), B(y_0)) \leqslant \frac{\varepsilon}{1-\theta};$$

en passant à la limite pour $n \to \infty$, on obtient

$$\rho(y_0, z_0) \leqslant \frac{\varepsilon}{1-\theta},$$

ce qu'il nous fallait.

§ 13.3. Existence et unicité de la solution d'une équation différentielle dans un espace normé

13.31. Soient B un espace de Banach et $\Phi(t, x)$ une application de l'espace B dans lui-même qui dépend d'un paramètre réel t, $a \leqslant t \leqslant b$. Soit $u(t)$ une fonction vectorielle dérivable définie sur le même intervalle $a \leqslant t \leqslant b$, à valeurs dans l'espace B. Nous savons que la dérivée $u'(t)$ est encore une fonction vectorielle définie sur le même intervalle $a \leqslant t \leqslant b$, à valeurs dans le même espace B. Si l'on substitue à la variable $x \in$ B, dans la fonction $\Phi(t, x)$, la fonction vectorielle $u(t)$, on obtient une nouvelle fonction vectorielle $\Phi[t, u(t)]$, à valeurs dans l'espace B, définie pour $t \in [a, b]$.

Nous tenons à résoudre l'équation différentielle

$$u'(t) = \Phi[t, u(t)] \tag{1}$$

avec la condition initiale

$$u(t_0) = u_0, \quad a < t_0 < b, \quad u_0 \in \text{B}. \tag{2}$$

Dans l'hypothèse naturelle de continuité qui sera précisée par la suite, la recherche d'une solution de l'équation (1) avec la condition initiale (2) est équivalente à la recherche d'une solution de l'équation intégrale

$$u(t) = u_0 + \int_{t_0}^{t} \Phi[\tau, u(\tau)]\, d\tau, \tag{3}$$

parce que (3) se déduit de (1) et (2) en intégrant de t_0 à t, (2) se déduit de (3) par la substitution $t = t_0$ et (1) s'obtient de (3) en dérivant

par rapport à t. Ainsi, le problème est ramené à la recherche d'un point fixe de l'application

$$A[x(t)] = u_0 + \int_{t_0}^{t} \Phi[\tau,\ x(t)]\, d\tau \qquad (4)$$

dans l'espace des fonctions vectorielles $x(t)$.

13.32. Naturellement, nous allons appliquer le principe du point fixe de Picard-Banach. Pour le faire, on doit disposer d'un espace métrique complet M et d'une application contractante A convenant à notre problème.

Choisissons pour M l'espace de toutes les fonctions vectorielles continues $x(t)$, à valeurs dans B, qui sont définies dans un intervalle $[t_0 - h,\ t_0 + h]$; la valeur de h sera indiquée plus bas. Munissons l'espace M de la métrique suivante

$$\rho[x_1(t),\ x_2(t)] = \max_{|t-t_0| \leqslant h} \| x_1(t) - x_2(t) \|.$$

L'espace métrique M ainsi obtenu est *complet* (12.23*f*).

13.33. L'application A de l'espace M dans lui-même doit être donnée par la formule 13.31 (4). Précisons les conditions à imposer à la fonction $\Phi(t, x)$ pour que la définition de l'application A soit correcte. Notamment, nous supposons que l'application $\Phi(t, x)$ soit continue par rapport à l'ensemble des variables t et x; cela veut dire que quels que soient t_1, x_1 et $\varepsilon > 0$, il existe un $\delta = \delta(t_1, x_1, \varepsilon)$ tel que $\| \Phi(t_1, x_1) - \Phi(t_2, x_2) \| < \varepsilon$ dès que $\| x_1 - x_2 \| < \delta$, $| t_1 - t_2 | < \delta$. Dans cette supposition, *la fonction vectorielle* $\Phi(t, x(t))$ *est continue en* t quelle que soit une fonction continue $x(t)$. En effet, pour t_1 $\varepsilon > 0$ donnés, posons $x(t_1) = x_1$ et trouvons un $\delta > 0$ de façon à avoir $\| \Phi(t_1, x_1) - \Phi(t_2, x_2) \| < \varepsilon$ dès que $\| x_1 - x_2 \| < \delta$, $| t_1 - t_2 | < \delta$. Puis trouvons un $\delta_1 > 0$ tel que $| t_1 - t_2 | < \delta_1$ implique l'inégalité $\| x(t_1) - x(t_2) \| < \delta$. Alors, pour $| t_1 - t_2 | < \min(\delta, \delta_1)$, nous avons bien $\| \Phi(t_1, x(t_1)) - \Phi(t_2, x(t_2)) \| < \varepsilon$.

13.34. Ensuite, pour que l'hypothèse du théorème du point fixe (A est une application contractante) soit remplie, nous supposons que la fonction $\Phi(t, x)$ satisfasse à *la condition de Lipschitz* que voici: il existe une constante C telle que

$$\| \Phi(t, x_1) - \Phi(t, x_2) \| \leqslant C \| x_1 - x_2 \| \qquad (1)$$

pour deux éléments quelconques x_1 et x_2 de l'espace B.

Montrons que, dans les conditions formulées, l'application 13.31(4) est contractante au moins pour h suffisamment petit. En effet, quels que soient deux points de l'espace M, i.e. deux fonctions vectorielles $x(t)$ et $y(t)$ définies sur $[t_0 - h,\ t_0 + h]$ et con-

tinues, nous avons

$$\rho(A[x(t)], A[y(t)]) = \max_{|t-t_0|\leqslant h} \| A[x(t)] - A[y(t)] \| =$$

$$= \max \left\| \int_{t_0}^{t} \{\Phi(\tau, x(\tau)) - \Phi(\tau, y(\tau))\} d\tau \right\| \leqslant$$

$$\leqslant h \max_{|t-t_0|\leqslant h} \| \Phi(t, x(t)) - \Phi(t, y(t)) \| \leqslant$$

$$\leqslant Ch \max_{|t-t_0|\leqslant h} \| x(t) - y(t) \| = Ch\rho(x, y), \qquad (2)$$

et, pour que l'application soit contractante, il suffit de choisir $h < 1/C$.

13.35. Nous sommes maintenant en droit d'appliquer le principe du point fixe pour démontrer *l'existence et l'unicité de la solution de l'équation* 13.31 (1) *pour la condition initiale* 13.31 (2). Pour le moment, cette solution n'est définie que sur $[t_0 - h, t_0 + h]$, mais il est possible, en appliquant successivement le résultat établi, prolonger-la sur tout l'intervalle $[a, b]$. A savoir, fixons la valeur $h = 2/(3C)$ et appliquons le théorème démontré à la même équation différentielle 13.31 (1), mais avec la condition initiale

$$t_1 = t_0 + h, \quad u^*(t_1) = u(t_1),$$

où $u(t_1)$ est la valeur, pour $t = t_1$, de la solution déjà construite définie sur $[t_0 - h, t_0 + h]$. Nous aboutissons à une nouvelle solution $u^*(t)$ définie sur l'intervalle $[t_1 - h, t_1 + h]$. Or, en vertu de l'unicité démontrée de la solution, les fonctions $u^*(t)$ et $u(t)$ se confondent sur la partie commune de leurs domaines de définition, et nous avons donc affaire à une solution de l'équation 13.31 (3) définie sur l'intervalle $[t_0 - h, t_0 + 2h]$. En continuant nous arrivons, après un nombre fini de pas, à une solution définie sur tout l'intervalle $[a, b]$.

Définitivement, nous avons démontré le théorème suivant:

Théorème. *Si la fonction* $\Phi(t, x)$ *est définie pour* $a \leqslant t \leqslant b$, $x \in B$, *continue par rapport à l'ensemble de ces variables et satisfait sur tout l'intervalle* $a \leqslant t \leqslant b$ *à la condition de Lipschitz* 13.34 (1), *alors l'équation* 13.31 (1) *avec la condition* 13.31 (2) *possède une solution* $u = u(t)$ *définie sur tout l'intervalle* $[a, b]$, *qui est unique dans la classe de toutes les fonctions vectorielles dérivables* $x(t)$, $t \in [a, b]$, *à valeurs dans l'espace* B.

13.36. Signalons le cas où la fonction vectorielle $\Phi(t, x)$, pour tout $t \in [a, b]$, applique un sous-espace fermé fixe $B_1 \subset B$ dans lui-même. Dans cette condition, si l'on choisit un vecteur initial u_0 dans le même sous-espace B_1, alors *la solution correspondante* $u(t)$

appartiendra au sous-espace B_1 pour tous les $t \in [a, b]$. En effet, dans cette situation, nous pouvons dès le début considérer le sous-espace B_1 au lieu de l'espace B. Nous aboutissons à une solution dans le sous-espace B_1; vu le théorème d'unicité, il n'existe dans l'espace B aucune autre solution de l'équation 13.31 (1) qui vérifie la condition $u(t_0) = u_0 \in B_1$.

13.37. Solution comme fonction continue du vecteur initial. Ici la solution de l'équation 13.31 (1) avec la condition initiale 13.31 (2) est désignée par $u(t; t_0, u_0)$. C'est une fonction qui fait correspondre à un vecteur $u_0 \in B$ le vecteur $u(t; t_0, u_0)$ qui dépend de t_0 et t comme des paramètres numériques. Montrons que, *dans l'hypothèse du théorème* 13.35, *pour t_0 et t fixes, le vecteur $u(t; t_0, u_0)$ est une fonction continue de u_0.*

Considérons les applications

$$A[x(t)] = u_0 + \int_{t_0}^{t} \Phi[\tau, x(\tau)]\, d\tau,$$

$$B[x(t)] = u_1 + \int_{t_0}^{t} \Phi[\tau, x(\tau)]\, d\tau$$

dont les points fixes sont les solutions de l'équation 13.31 (1) avec u_0 et u_1 pour vecteurs initiaux respectifs. Dans l'espace M des fonctions vectorielles $x(t)$ définies et continues sur $[t_0 - h, t_0 + h]$, où $h < 1/C$, ce sont des applications contractantes avec une même valeur $\theta = Ch < 1$. Si $|u_0 - u_1| < \varepsilon$, ces applications sont ε-proches (13.23), donc la distance de leurs points fixes ne dépasse pas $\varepsilon/(1 - \theta)$. Autrement dit,

$$\max_{|t - t_0| \leqslant h} |u(t; t_0, u_0) - u(t; t_0, u_1)| < \frac{\varepsilon}{1 - \theta}.$$

Ainsi, lorsque la différence des deux solutions, pour $t = t_0$, est inférieure à ε, elle est inférieure à $\varepsilon/(1 - \theta)$ dans l'intervalle $|t - t_0| \leqslant h$. En transférant le point initial de t_0 à $t_1 = t_0 + h$ nous obtenons, comme dans 13.35, une possibilité de prolonger la solution dans l'intervalle $t_0 \leqslant t_0 + h \leqslant t_0 + 2h$; en répétant le procédé nous voyons que l'écart des solutions dans cet intervalle ne dépasse pas $\varepsilon/(1 - \theta)^2$. En continuant nous arrivons à l'estimation

$$\max_{a \leqslant t \leqslant b} |u(t; t_0, u_0) - u(t; t_0, u_1)| < \frac{\varepsilon}{(1 - \theta)^m},$$

où $m = \left[\frac{b - a}{h}\right] + 1$; notre proposition est donc démontrée.

13.38. Opérateur résolvant. Soient $u_0 \in B$ un vecteur quelconque et $u(t)$ une solution de l'équation 13.31 (1) avec

$u(t_0) = u_0$ pour condition initiale. Pour un $t \in [a, b]$ fixe, le vecteur $u = u(t)$ est défini d'une manière univoque. Il dépend de u_0 ainsi que de t_0 et t. Notons ce fait comme suit :

$$u(t) = \Omega_{t_0}^{t}(u_0).$$

Ici $\Omega_{t_0}^{t}$ est une application de l'espace B dans lui-même que nous appellerons *opérateur résolvant de l'équation* 13.31 (1).

Ainsi, pour *l'équation linéaire homogène*

$$u'(t) = Au(t)$$

l'opérateur résolvant est de la forme (13.13)

$$\Omega_{t_0}^{t} = e^{(t-t_0)A}.$$

a. En vertu de 13.37, l'opérateur $\Omega_{t_0}^{t}$ est *continu* : si une suite $u_0^{(1)}, u_0^{(2)}, \ldots, u_0^{(n)}, \ldots$ de vecteurs de l'espace B tend vers un vecteur u_0, alors la suite correspondante $u_1^{(n)} = \Omega_{t_0}^{t}(u_0^{(n)})$ tend vers le vecteur $u_1 = \Omega_{t_0}^{t}(u_0)$.

b. Il est évident que $\Omega_{t_0}^{t_0}(u_0) = u_0$, de sorte que $\Omega_t^t = E$ est un opérateur identique.

c. Démontrons l'égalité

$$\Omega_{t_0}^{t_2} = \Omega_{t_1}^{t_2}\Omega_{t_0}^{t_1} \tag{1}$$

quels que soient t_0, t_1, t_2 de $[a, b]$.

En effet, soient $u_1 = \Omega_{t_0}^{t_1}(u_0)$ et $u_2 = \Omega_{t_1}^{t_2}(u_1)$. Le vecteur u_1 est la valeur, pour $t = t_1$, de la solution $u(t)$ de l'équation 13.31 (1) qui prend la valeur u_0 pour $t = t_0$. Le vecteur u_2 est la valeur, pour $t = t_2$, de la solution qui prend la valeur u_1 pour $t = t_1$. Vu l'unicité, ces deux solutions se confondent, ce qui démontre (1).

d. En posant dans (1) $t_2 = t_0$ on a $E = \Omega_{t_1}^{t_0}\Omega_{t_0}^{t_1}$, d'où l'inversibilité de l'opérateur $\Omega_{t_0}^{t_1}$.

e. L'égalité $\frac{du(t)}{dt} = A(t)u(t)$ peut être mise sous la forme

$$\frac{d(\Omega_{t_0}^{t}(u_0))}{dt} = A(t)[\Omega_{t_0}^{t}(u_0)], \tag{2}$$

ou

$$\frac{d\Omega_{t_0}^{t}}{dt} = A(t)\Omega_{t_0}^{t} \tag{3}$$

en sous-entendant par la dernière notation que l'égalité (2) a lieu pour tout élément $u_0 \in B$.

13.39. Nous avons supposé que la fonction $\Phi(t, x)$ soit définie pour tous les $t \in [a, b]$ et $x \in B$. L'existence et l'unicité de la solution de l'équation 13.31 (1) dans un voisinage d'un point t_0 avec

la condition 13.31 (2) peuvent être démontrées dans une hypothèse plus faible, à savoir: la fonction $\Phi(t, x)$ doit être définie pour $a \leqslant t \leqslant b$ et ne doit être continue et satisfaire à la condition de Lipschitz que dans une boule

$$V = \{x \in \mathrm{B} : \| x - u_0 \| \leqslant r\}.$$

Dans ce cas, si h est suffisamment petit, l'opérateur A transforme toute fonction continue $x(t)$, à valeurs dans la boule V, en une fonction de la même boule. Par conséquent, on peut réaliser la démonstration de l'existence et de l'unicité de la solution en remplaçant l'espace métrique M de *toutes* les fonctions $x(t)$ continues sur $[t_0 - h, t_0 + h]$ par l'espace métrique M_V des fonctions $x(t)$ *à valeurs dans la boule V*. Cependant, la solution ainsi obtenue ne sera en général pas prolongeable sur tout l'intervalle $a \leqslant t \leqslant b$.

Une situation analogue a lieu dans le cas où la fonction $\Phi(t, x)$ est définie et continue, pour $a \leqslant t \leqslant b$, dans tout l'espace B, mais la constante C de la condition de Lipschitz dépend des distances de l'origine des coordonnées aux points x_1 et x_2, de sorte que la condition de Lipschitz prend la forme

$$\| \Phi(t, x_1) - \Phi(t, x_2) \| \leqslant C(r) \| x_1 - x_2 \|$$

quels que soient x_1 et x_2 de la boule $\| x \| \leqslant r$. Dans ce cas, comme ci-dessus, la solution existe et est unique dans un voisinage de la valeur $t_0 \in [a, b]$, mais elle n'est en général pas prolongeable jusqu'à la frontière de l'intervalle $[a, b]$.

En tant qu'exemple, considérons l'équation $x'(t) = x^2(t)$ sur l'intervalle $-1 \leqslant t \leqslant 1$. Son second membre est continu pour tout $x \in R_1$; ensuite on a

$$| x_1^2 - x_2^2 | = | x_1 + x_2 | \, | x_1 - x_2 | \leqslant 2r | x_1 - x_2 |$$

quels que soient x_1 et x_2 de l'intervalle $| x | \leqslant r$. La solution de l'équation donnée, pour la condition initiale $x(0) = x_0$, a la forme

$$x(t) = \frac{x_0}{1 - t x_0}$$

et n'est pas prolongeable sur tout l'intervalle $-1 \leqslant t \leqslant 1$ si $| x_0 | \geqslant 1$.

§ 13.4. Système d'équations vectorielles

13.41. Soit toujours B un espace de Banach et soit n fonctions

$$\Phi_1(t, x_1, \ldots, x_n), \ldots, \Phi_n(t, x_1, \ldots, x_n)$$

dont chacune dépend d'un paramètre réel $t \in [a, b]$ et de n variables $x_1, \ldots, x_n$ parcourant B, chaque fonction $\Phi_k(t, x_1, \ldots, x_n)$ prenant ses valeurs toujours dans l'espace B. Considérons

le système d'équations différentielles

$$\left.\begin{aligned} u_1'(t) &= \Phi_1(t, u_1, \ldots, u_n), \\ u_2'(t) &= \Phi_2(t, u_1, \ldots, u_n), \\ &\cdots\cdots\cdots\cdots \\ u_n'(t) &= \Phi_n(t, u_1, \ldots, u_n) \end{aligned}\right\} \qquad (1)$$

avec les conditions initiales

$$u_1(t_0) = p_1 \in \mathrm{B}, \ldots, u_n(t_0) = p_n \in \mathrm{B}, \ a \leqslant t_0 \leqslant b. \qquad (2)$$

Naturellement, on appelle *solution du système* (1) *avec les conditions* (2) *un système* de fonctions vectorielles $u_1(t), \ldots, u_n(t)$ définies pour $a \leqslant t \leqslant b$, vérifiant toutes les équations du système (1) ainsi que les conditions (2).

Par définition, une fonction $\Phi_k(t, x_1, \ldots, x_n)$ est *continue par rapport à l'ensemble des variables* $t, x_1, \ldots, x_n$ si, quels que soient $t, x_1, \ldots, x_n$ et $\varepsilon > 0$, il existe un $\delta > 0$ tel que les relations

$$|\bar{t} - t_0| < \delta, \quad \|\bar{x}_1 - x_1\| < \delta, \ldots, \|\bar{x}_n - x_n\| < \delta$$

impliquent l'inégalité

$$\|\Phi_k(\bar{t}, \bar{x}_1, \ldots, \bar{x}_n) - \Phi_k(t, x_1, \ldots, x_n)\| \leqslant \varepsilon.$$

Par définition, la fonction $\Phi_k(t, x_1, \ldots, x_n)$ satisfait à la *condition de Lipschitz* par rapport aux variables $x_1, \ldots, x_n$ s'il existe une constante C telle que

$$\|\Phi_k(t, \bar{x}_1, \ldots, \bar{x}_n) - \Phi_k(t, x_1, \ldots, x_n)\| \leqslant$$

$$\leqslant C \sum_{j=1}^{n} \|\bar{x}_j - x_j\|$$

quels que soient $2n$ éléments $x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n$ de l'espace B.

13.42. Le théorème suivant a lieu:

Théorème. *Si toutes les fonctions* $\Phi_k(t, x_1, \ldots, x_n)$ *sont continues par rapport à l'ensemble des variables* $t, x_1, \ldots, x_n$ *et satisfont à la condition de Lipschitz par rapport aux variables* $x_1, \ldots, x_n$, *alors le système* 13.41 (1) *avec les conditions* 13.41 (2) *possède une solution* $(u_1(t), \ldots, u_n(t))$ *et une seule dans la classe de tous les complexes* $(x_1(t), \ldots, x_n(t))$ *que l'on peut construire avec les fonctions vectorielles dérivables à valeurs dans l'espace* B.

Démonstration. Construisons un nouvel espace vectoriel normé B^n des complexes $x = (x_1, \ldots, x_n)$ formés chacun de n éléments de l'espace B. Les opérations linéaires dans l'espace B^n s'effectuent par coordonnées: si $x = (x_1, \ldots, x_n) \in \mathrm{B}^n$,

$y = (y_1, \ldots, y_n) \in B^n$, alors

$$x + y = (x_1 + y_1, \ldots, x_n + y_n),$$
$$\alpha x = (\alpha x_1, \ldots, \alpha x_n).$$

Les propriétés nécessaires des opérations linéaires ainsi introduites s'établissent facilement à partir des propriétés respectives des opérations linéaires dans l'espace B. Ensuite, nous choisissons pour norme dans B^n la quantité

$$|||x||| = \sum_{j=1}^{n} \|x_j\|. \tag{1}$$

Les propriétés nécessaires de la norme dans l'espace B^n se déduisent facilement des propriétés respectives de la norme dans l'espace B. La convergence en norme (1) dans l'espace B^n est la convergence par rapport à chaque coordonnée dans l'espace B. Enfin, l'espace B étant complet, on démontre aisément la même propriété pour B^n.

On peut considérer le système de fonctions

$$\left.\begin{aligned} y_1 &= \Phi_1(t, x_1, \ldots, x_n), \\ y_2 &= \Phi_2(t, x_1, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots \\ y_n &= \Phi_n(t, x_1, \ldots, x_n) \end{aligned}\right\} \tag{2}$$

comme une seule application $y = \Phi(t, x)$ de l'espace B^n dans lui-même. Montrons que, dans les hypothèses sur les fonctions $\Phi_k(t, x_1, \ldots, x_n)$ formulées plus haut, la fonction $\Phi(t, x)$ est continue par rapport à l'ensemble des variables et vérifie la condition de Lipschitz par rapport à la variable x. Pour $t, x_1, \ldots, x_n$ et $\varepsilon > 0$ donnés, trouvons un nombre δ à partir de la condition de continuité de toutes les fonctions $\Phi_k(t, x_1, \ldots, x_n)$ de sorte que si $|\bar{t} - t| < \delta$ et si $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$ est tel que

$$|||\bar{x} - x||| = \sum_{j=1}^{n} \|\bar{x}_j - x_j\| < \delta,$$

alors on a

$$\|\Phi_j(\bar{t}, \bar{x}_1, \ldots, \bar{x}_n) - \Phi_j(t, x_1, \ldots, x_n)\| < \frac{\varepsilon}{n}.$$

Il en résulte que

$$|||\Phi(\bar{t}, \bar{x}) - \Phi(t, x)||| =$$
$$= \sum_{j=1}^{n} \|\Phi_j(\bar{t}, \bar{x}_1, \ldots, \bar{x}_n) - \Phi_j(t, x_1, \ldots, x_n)\| \leqslant \varepsilon;$$

par conséquent, la fonction $\Phi(t, x)$ est continue par rapport à l'ensemble des variables. Ensuite, il résulte de la condition de Lipschitz

pour les fonctions $\Phi_k(t, x_1, \ldots, x_n)$ que

$$|||\Phi(t, \bar{x}) - \Phi(t, x)||| =$$

$$= \sum_{j=1}^{n} \|\Phi_j(t, \bar{x}_1, \ldots, \bar{x}_n) - \Phi_j(t, x_1, \ldots, x_n)\| \leqslant$$

$$\leqslant \sum_{j=1}^{n} \sum_{k=1}^{n} C \|\bar{x}_k - x_k\| = Cn |||\bar{x} - x|||,$$

de sorte que $\Phi(t, x)$ satisfait à la condition de Lipschitz par rapport à la variable x.

D'après le théorème 13.35, l'équation différentielle

$$u'(t) = \Phi(t, x) \tag{3}$$

pour une fonction vectorielle $u(t)$ à valeurs dans l'espace B^n, avec la condition initiale

$$u(t_0) = p = (p_1, \ldots, p_n) \in B^n \tag{4}$$

possède une solution $u(t)$ définie pour $t \in [a, b]$ qui est unique dans la classe de toutes les fonctions vectorielles dérivables $x(t)$ à valeurs dans l'espace B^n. Etant donné que, conformément aux définitions,

$$u(t) = (u_1(t), \ldots, u_n(t)), \quad u'(t) = (u_1'(t), \ldots, u_n'(t)),$$

l'équation (3) avec la condition (4) pour une fonction $u(t)$ est équivalente au système 13.41 (1) avec les conditions 13.41 (2) pour les fonctions $u_1(t), \ldots, u_n(t)$. Le théorème est démontré.

13.43. S'il existe un sous-espace fermé $B_1 \subset B$ tel que, quels que soient $t \in [a, b]$ et $x_1, x_2, \ldots, x_n \in B_1$, les valeurs $\Phi_k(t, x_1, \ldots, x_n)$ $(k = 1, \ldots, n)$ et les vecteurs initiaux $p_1, \ldots, p_n$ appartiennent à B_1, alors les valeurs de toutes les fonctions $u_1(t), \ldots, u_n(t)$ représentant la solution du système 13.41 (1) avec les conditions 13.41 (2) appartiennent elles aussi à l'espace B_1 pour tout $t \in [a, b]$.

En effet, considérons dans B^n le sous-espace B_1^n des vecteurs $x = (x_1, \ldots, x_n)$ dont chaque coordonnée appartient au sous-espace $B_1 \subset B$. Il est aisé de prouver que B_1^n est fermé dans B^n. Par hypothèse, la transformation 13.42 (2) applique B_1^n dans lui-même. Par conséquent, compte tenu de la remarque 13.39, si le vecteur $(p_1, \ldots, p_n)$ est de B_1^n, il en est de même de la solution $(u_1(t), \ldots, u_n(t))$, pour tout $t \in [a, b]$, ce qu'il fallait démontrer.

§ 13.5. Equation vectorielle d'ordre supérieur

13.51. Considérons encore une fois un espace de Banach B et soit $\Phi(t, x_1, \ldots, x_n)$ une fonction d'un paramètre réel t, $a \leqslant t \leqslant b$, et de n points $x_1, \ldots, x_n$ de l'espace B, à valeurs dans le

même espace B. Soit

$$u^{(m)}(t) = \Phi(t, u(t), \ldots, u^{(m-1)}(t)) \tag{1}$$

une équation différentielle d'ordre m complétée par les conditions initiales

$$u(t_0) = p_1 \in \mathrm{B}, \quad u'(t_0) = p_2 \in \mathrm{B}, \ldots, u^{(m-1)}(t_0) = p_m \in \mathrm{B}. \tag{2}$$

Théorème. *Si la fonction $\Phi(t, x_1, \ldots, x_n)$ est continue par rapport à l'ensemble des variables $t, x_1, \ldots, x_m$ et satisfait à la condition de Lipschitz par rapport aux variables $x_1, \ldots, x_m$, alors l'équation (1) avec les conditions (2) possède une solution $u = u(t)$ et une seule dans la classe de toutes les fonctions m fois dérivables $x(t)$ à valeurs dans* B.

Démonstration. En plus de l'équation (1) et des conditions (2), considérons le système d'équations différentielles

$$\left.\begin{aligned} u_1'(t) &= u_2(t), \\ u_2'(t) &= u_3(t), \\ &\cdots\cdots\cdots \\ u_{m-1}'(t) &= u_m(t), \\ u_m'(t) &= \Phi(t, u_1(t), \ldots, u_m(t)) \end{aligned}\right\} \tag{3}$$

avec les conditions initiales

$$u_1(t_0) = p_1, \ldots, u_m(t_0) = p_m. \tag{4}$$

Le système (3) est un cas particulier du système 13.41 (1) que l'on obtient en posant

$$\left.\begin{aligned} \Phi_1(t, x_1, \ldots, x_m) &\equiv x_2, \\ \Phi_2(t, x_1, \ldots, x_m) &\equiv x_3, \\ &\cdots\cdots\cdots \\ \Phi_{m-1}(t, x_1, \ldots, x_m) &\equiv x_m, \\ \Phi_m(t, x_1, \ldots, x_m) &\equiv \Phi(t, x_1, \ldots, x_m). \end{aligned}\right\} \tag{5}$$

Dans le présent cas, toutes les fonctions

$$\Phi_k(t, x_1, \ldots, x_m) \quad (k = 1, \ldots, m)$$

sont continues par rapport à l'ensemble des variables $t, x_1, \ldots, x_m$ et satisfont à la condition de Lipschitz par rapport aux variables $x_1, \ldots, x_m$; c'est évident pour les $m - 1$ premières fonctions Φ_k et donné par hypothèse pour la dernière.

Par conséquent, en vertu du théorème 13.42, le système (3) avec les conditions (4) possède une solution $u_1(t), \ldots, u_m(t)$. Posons $u(t) = u_1(t)$. La première équation du système (3) montre

que $u'(t) = u_2(t)$, la suivante que $u''(t) = u_2'(t) = u_3(t)$, etc.; la $(m-1)$-ième équation montre que $u^{(m-1)}(t) = u_{m-1}'(t) = u_m(t)$, et enfin la dernière que

$$u^{(m)}(t) = u_m'(t) = \Phi(t, u, u', \ldots, u^{(m-1)}).$$

Ainsi, la fonction vectorielle $u(t)$ vérifie l'équation (1). Comme les conditions (4) sont également satisfaites, cette fonction vérifie les conditions (2). De la sorte, l'équation (1) avec les conditions (2) *possède une solution*. Montrons qu'elle est unique. Si $\bar{u}(t)$ est une solution quelconque de l'équation (1) avec les conditions (2), alors le système de fonctions

$$\bar{u}_1(t) \equiv \bar{u}(t), \quad \bar{u}_2(t) \equiv \bar{u}'(t), \ldots, \bar{u}_m(t) = \bar{u}^{(m-1)}(t)$$

vérifie, évidemment, le système (3) avec les conditions (4); comme, d'après le théorème 13.42, la solution du système (3) avec les conditions (4) est unique, nous avons $\bar{u}(t) \equiv \bar{u}_1(t) \equiv u_1(t) \equiv u(t)$, ce qu'il fallait démontrer.

13.52. Supposons qu'il existe un sous-espace fermé $B_1 \subset B$ tel que la fonction $\Phi(t, x_1, \ldots, x_m)$ prend ses valeurs dans B_1 quels que soient $t \in [a, b]$ et $x_1 \in B_1, \ldots, x_m \in B_1$.

Si, de plus, les vecteurs $p_1, \ldots, p_m$ qui figurent dans les conditions initiales pour l'équation (1) appartiennent eux aussi au sous-espace B_1, alors la solution correspondante $u(t)$ appartient également à B_1 pour tout $t \in [a, b]$.

En effet, dans les conditions formulées, toutes les fonctions du système 13.51 (5) ont leurs valeurs dans B_1 si $x_1 \in B_1, \ldots, x_n \in B_1$. Il résulte de $p_1 \in B_1, \ldots, p_m \in B_1$, en vertu de la remarque 12.53, que $u_1(t) \in B_1, \ldots, u_m(t) \in B_1$ pour tout $t \in [a, b]$. Comme $u_1(t) \equiv u(t)$, nous arrivons au résultat cherché.

§ 13.6. Equations et systèmes linéaires

13.61. Considérons un opérateur linéaire borné $A(t)$ dans un espace vectoriel normé B, qui dépend d'un paramètre t, $a \leqslant t \leqslant b$. Nous disons que l'opérateur $A(t)$ *dépend continûment de* t si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $\| A(\bar{t}) - A(t) \| < \varepsilon$ dès que $|\bar{t} - t| < \delta$ (ici $\| \ \|$ désigne la *norme d'un opérateur linéaire* (12.71*b*)).

Une fonction $\Phi(t, x)$ à valeurs dans l'espace B est dite *du premier degré par rapport à la variable* $x \in B$ si

$$\Phi(t, x) = A(t)\, x + b(t), \tag{1}$$

où $A(t)$ est un opérateur linéaire borné qui dépend continûment du paramètre t et $b(t)$ une fonction continue à valeurs dans l'espace B.

Montrons qu'une fonction $\Phi(t, x)$ du type (1) *est continue par rapport à l'ensemble des variables* t, x.

L'opérateur $A(t)$ peut être considéré comme fonction continue de t à valeurs dans l'espace normé $L(B)$ de tous les opérateurs linéaires bornés dans B (12.73*a*). Une telle fonction est bornée sur l'intervalle $a \leqslant t \leqslant b$, donc

$$\sup_{a \leqslant t \leqslant b} \| A(t) \| \equiv A < \infty.$$

Pour $\varepsilon > 0$, t et x donnés, trouvons un $\delta = \delta(\varepsilon, t, x)$ de façon à avoir les inégalités

$$\| x \| \| A(\bar{t}) - A(t) \| \leqslant \frac{\varepsilon}{3},$$

$$\| b(t) - b(t) \| \leqslant \frac{\varepsilon}{3}$$

pour $|\bar{t} - t| < \delta$. Alors, pour les mêmes $\bar{t}$, t et pour $\| \bar{x} - x \| < \varepsilon/(3A)$, nous avons

$$\| \Phi(\bar{t}, \bar{x}) - \Phi(t, x) \| =$$

$$= \| A(\bar{t}) \bar{x} - A(\bar{t}) x + A(\bar{t}) x - A(t) x + b(\bar{t}) - b(t) \| \leqslant$$

$$\leqslant \| A(\bar{t}) \| \| \bar{x} - x \| + \| A(\bar{t}) - A(t) \| \| x \| + \| b(\bar{t}) - b(t) \| \leqslant$$

$$\leqslant A \frac{\varepsilon}{3A} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

et la continuité demandée est établie.

Montrons à présent qu'une fonction du type (1) *satisfait à la condition de Lipschitz par rapport à la variable* x, *avec la constante* $C = A = \sup_{a \leqslant t \leqslant b} \| A(t) \|$.

En effet, d'après la définition de la norme d'un opérateur, nous avons

$$\| \Phi(t, \bar{x}) - \Phi(t, x) \| = \| A(t) \bar{x} - A(t) x \| \leqslant$$

$$\leqslant \| A(t) \| \| \bar{x} - x \| \leqslant A \| \bar{x} - x \|,$$

ce qu'il fallait démontrer.

13.62. En appliquant le théorème 13.35 nous arrivons au résultat suivant :

Théorème. *Une équation différentielle linéaire*

$$u'(t) = A(t) u(t) + b(t) \tag{1}$$

où $A(t)$ *est un opérateur linéaire borné dans l'espace* B, *qui dépend continûment du paramètre* $t \in [a, b]$ *et* $b(t)$ *est une fonction continue*

à valeurs dans B, *avec la condition initiale*

$$u(t_0) = u_0, \tag{2}$$

possède une solution $u = u(t)$ *et une seule qui est une fonction dérivable à valeurs dans* B.

13.63. Systèmes d'équations linéaires. Soit un système d'équations linéaires

$$\left.\begin{aligned} u_1'(t) &= A_{11}(t)\,u_1(t) + \ldots + A_{1n}(t)\,u_n(t) + b_1(t), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ u_n'(t) &= A_{n1}(t)\,u_1(t) + \ldots + A_{nn}(t)\,u_n(t) + b_n(t), \end{aligned}\right\} \tag{1}$$

où $A_{jk}(t)$ $(j, k = 1, \ldots, n)$ sont des opérateurs linéaires bornés dans l'espace B qui dépendent continûment du paramètre $t \in [a, b]$ et $b_1(t), \ldots, b_n(t)$ des fonctions vectorielles continues de t à valeurs dans B. Le système (1) est complété par la condition initiale

$$u_1(t_0) = p_1 \in B, \ldots, u_n(t_0) = p_n \in B, \quad a \leqslant t_0 \leqslant b. \tag{2}$$

Théorème. *Le système* (1) *avec la condition initiale* (2) *a une solution* $(u_1(t), \ldots, u_n(t))$ *composée des fonctions vectorielles de* $t \in [a, b]$, *à valeurs dans* B *et une seule.*

Démonstration. Le système (1) est un cas particulier du système considéré dans 13.41:

$$\begin{aligned} u_1'(t) &= \Phi_1(t, u_1, \ldots, u_n), \\ &\ldots\ldots\ldots\ldots \\ u_n'(t) &= \Phi_n(t, u_1, \ldots, u_n). \end{aligned}$$

Ce cas s'obtient en posant

$$\Phi_k(t, x_1, \ldots, x_n) = A_{k1}(t)\,x_1 + \ldots + A_{kn}(t)\,x_n + b_k(t), \quad k = 1, \ldots, n. \tag{3}$$

Pour appliquer le théorème 13.42, il faut prouver que toute fonction (3) est continue par rapport à l'ensemble des variables t, $x_1, \ldots, x_n$ et satisfait à la condition de Lipschitz par rapport aux variables $x_1, \ldots, x_n$. Or, nous avons vu dans 13.61 que chaque terme $A_{km}(t)\,x_m$ et $b_k(t)$ vérifie ces conditions; il en est donc de même de leur somme (3). L'application du théorème 13.42 achève notre démonstration.

13.64. Si les opérateurs A_{jk} $(j, k = 1, \ldots, n)$, pour tout $t \in [a, b]$, appliquent un sous-espace fixe $B_1 \subset B$ dans lui-même et si les fonctions $b_j(t)$, pour $t \in [a, b]$, prennent leurs valeurs dans ce sous-espace B_1, alors, quels que soient les vecteurs initiaux $p_1, \ldots, p_n$ du sous-espace B_1, toutes les fonctions $u_1(t), \ldots, u_n(t)$ formant la solution du système 13.63 (1) avec les conditions 13.63 (2) prennent leurs valeurs dans le sous-espace B_1.

En effet, dans les conditions formulées, les fonctions

$$\Phi_j(t, x_1, \ldots, x_n) \equiv A_{j1}(t)\, x_1 + \ldots + A_{jn}(t)\, x_n \quad (j=1, \ldots, n),$$

pour $x_1 \in B_1, \ldots, x_n \in B_1$, prennent leurs valeurs dans le sous-espace B_1, et l'on peut appliquer 13.43.

13.65. Equation linéaire d'ordre supérieur. Considérons une équation linéaire d'ordre n

$$u^{(n)}(t) = A_1(t)\, u(t) + \ldots + A_n(t)\, u^{(n-1)}(t) + b(t) \qquad (1)$$

par rapport à la fonction inconnue $u(t)$ à valeurs dans l'espace B, avec les conditions initiales

$$u(t_0) = p_1 \in B, \ldots, \; u^{(n-1)}(t_0) = p_n \in B. \qquad (2)$$

Ici $A_k(t)$, pour tout $t \in [a, b]$, est un opérateur linéaire borné dans l'espace B et $b(t)$ une fonction continue à valeurs dans le même espace.

Théorème. *L'équation linéaire* (1) *avec les conditions initiales* (2) *possède une solution* $u(t)$ *et une seule dans la classe de toutes les fonctions vectorielles* n *fois dérivables à valeurs dans l'espace* B.

Démonstration. L'équation (1) est un cas particulier de l'équation considérée dans 13.51 :

$$u^{(n)}(t) = \Phi(t, u, u', \ldots, u^{(n-1)}).$$

Ce cas s'obtient en posant

$$\Phi(t, x_1, \ldots, x_n) = A_1(t)\, x_1 + \ldots + A_n(t)\, x_n + b(t).$$

Nous avons vu dans 13.63 que la fonction Φ est continue par rapport à l'ensemble des variables $t, x_1, \ldots, x_n$ et vérifie la condition de Lipschitz par rapport à $x_1, \ldots, x_n$. Par conséquent, le théorème 13.51 est valable. En l'appliquant on obtient le résultat cherché.

13.66. Si les opérateurs $A_1(t), \ldots, A_n(t)$, pour tout $t \in [a, b]$, appliquent un sous-espace fixe $B_1 \subset B$ dans lui-même et si la fonction $b(t)$ prend toutes ses valeurs dans ce sous-espace, alors, quels que soient les vecteurs initiaux $p_1, \ldots, p_n$ de B_1, la solution $u(t)$ de l'équation 13.65 (1) avec les conditions 13.65 (2) appartient, pour tout $t \in [a, b]$, au sous-espace B_1.

En effet, dans les conditions formulées, l'hypothèse de la remarque 13.64 est vérifiée ; en l'appliquant nous obtenons, en particulier, que $u_1(t) \equiv u(t)$ appartient à B_1 pour tout $t \in [a, b]$, ce qu'il nous fallait.

§ 13.7. Opérateur résolvant d'une équation linéaire homogène

13.71. L'équation 13.62 (1) avec $b(t) \equiv 0$:

$$u'(t) = A(t)\,u(t) \tag{1}$$

s'appelle *équation linéaire homogène.*

L'équation homogène possède une solution évidente $u(t) \equiv 0$. Toute autre solution de l'équation homogène ne s'annule pour aucun $t \in [a, b]$ d'après le théorème d'unicité 13.62.

En additionnant et en multipliant par les nombres les solutions de l'équation homogène (1) on obtient d'autres solutions de la même équation. En effet, si $u_1(t)$ et $u_2(t)$ sont des solutions de l'équation (1), alors, quels que soient les nombres α_1 et α_2, on a

$$\begin{aligned}(\alpha_1 u_1(t) + \alpha_2 u_2(t))' &= \alpha_1 u_1'(t) + \alpha_2 u_2'(t) = \\ &= \alpha_1 A(t)\,u_1(t) + \alpha_2 A(t)\,u_2(t) = \\ &= A(t)\,[\alpha_1 u_1(t) + \alpha_2 u_2(t)],\end{aligned}$$

donc $\alpha_1 u_1(t) + \alpha_2 u_2(t)$ est encore une solution de l'équation (1). Considérons l'opérateur résolvant $\Omega_{t_0}^t$ (13.38) de l'équation homogène (1). Cet opérateur est linéaire, c'est-à-dire que, quels que soient deux vecteurs u_1, u_2 et deux nombres α_1, α_2, on a

$$\Omega_{t_0}^t[\alpha_1 u_1 + \alpha_2 u_2] = \alpha_1\,\Omega_{t_0}^t(u_1) + \alpha_2 \Omega_{t_0}^t(u_2).$$

En effet, le second membre considéré comme fonction de t est une combinaison linéaire de solutions de l'équation (1) dont la valeur pour $t = t_0$ est $\alpha_1 u_1 + \alpha_2 u_2$. D'après ce qu'on a vu, il représente une solution de l'équation (1). Le premier membre, d'après la définition même, est la solution de l'équation (1) qui vaut $\alpha_1 u_1 + \alpha_2 u_2$ pour $t = t_0$. En vertu du théorème d'unicité, ces solutions se confondent pour tout $t \in [a, b]$, ce qu'on affirmait.

Donc, pour l'équation linéaire homogène (1), l'opérateur résolvant $\Omega_{t_0}^t$ est un opérateur linéaire. Rappelons que, d'après 13.38, il est continu et inversible. Dans ce qui suit nous écrirons $\Omega_{t_0}^t u$ au lieu de $\Omega_{t_0}^t(u)$.

13.72. Etudions la structure de l'opérateur résolvant de l'équation homogène 13.71 (1) dans l'espace n-dimensionnel R_n des vecteurs $x = (\xi_1, \ldots, \xi_n)$.

Choisissons dans l'espace R_n n vecteurs linéairement indépendants quelconques $f_1, \ldots, f_n$. Alors, à l'équation vectorielle 13.71 (1) par rapport à la fonction vectorielle inconnue $u(t) = \sum u_k(t) f_k$

il correspond le système d'équations scalaires:

$$\left.\begin{aligned} u_1'(t) &= a_{11}(t)\,u_1(t) + \ldots + a_{1n}(t)\,u_n(t), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ u_n'(t) &= a_{n1}(t)\,u_1(t) + \ldots + a_{nn}(t)\,u_n(t). \end{aligned}\right\} \tag{1}$$

A l'opérateur résolvant $\Omega_{t_0}^t$ nous pouvons faire correspondre, d'après les règles générales 12.17*a*, la matrice dont la k-ième colonne est formée des coordonnées du vecteur $\Omega_{t_0}^t f_k$ ($k = 1, \ldots, n$). En d'autres termes, à l'opérateur résolvant $\Omega_{t_0}^t$, il correspond la matrice

$$W_{t_0}^t = \left\| \begin{matrix} f_{11}(t) & f_{12}(t) \ldots f_{1n}(t) \\ f_{21}(t) & f_{22}(t) \ldots f_{2n}(t) \\ \cdots\cdots\cdots\cdots \\ f_{n1}(t) & f_{n2}(t) \ldots f_{nn}(t) \end{matrix} \right\|, \tag{2}$$

où $f_{1k}(t), \ldots, f_{nk}(t)$ sont les coordonnées de la solution $f_k(t)$ dont la valeur pour $t = t_0$ est le vecteur f_k. La matrice (2) s'appelle *matrice de Wronski* du système (1), et son déterminant est le *déterminant de Wronski*, ou *wronskien* du système (1). Pour une matrice constante $A = \| a_{jk} \|$, nous avons écrit la matrice de Wronski dans 13.15.

L'opérateur $\Omega_{t_0}^t$ étant inversible, la matrice (2) est non dégénérée pour tout $t \in [a, b]$, et le déterminant de Wronski ne s'annule pour aucun $t \in [a, b]$; ainsi, les solutions $f_1(t), \ldots, f_n(t)$ qui sont linéairement indépendantes pour $t = t_0$ le restent pour tout $t \in [a, b]$.

La solution de l'équation 13.71 (1) avec un vecteur initial quelconque (pour $t = t_0$) $u = \sum_{k=1}^{n} u_k f_k$ est construite d'après la formule générale

$$u(t) = \Omega_{t_0}^t u = \Omega_{t_0}^t \sum_{k=1}^{n} u_k f_k = \sum_{k=1}^{n} u_k \Omega_{t_0}^t f_k ;$$

par conséquent, toute solution du système (1) est une combinaison linéaire de n solutions particulières $f_1(t), \ldots, f_n(t)$.

Nous voyons que, dans le présent cas, l'espace de toutes les solutions du système (1) est de dimension n et que les fonctions vectorielles $f_1(t), \ldots, f_n(t)$ en constituent une base. L'ensemble des fonctions vectorielles $f_1(t), \ldots, f_n(t)$ est appelé *système fondamental de solutions du système* (1).

13.73: Calcul du déterminant de Wronski. Le déterminant de Wronski

$$[f_1(t), \ldots, f_n(t)] = \begin{vmatrix} f_{11}(t) & f_{12}(t) \ldots f_{1n}(t) \\ \cdots\cdots\cdots\cdots \\ f_{n1}(t) & f_{n2}(t) \ldots f_{nn}(t) \end{vmatrix} \tag{1}$$

peut être appelé, par analogie avec l'algèbre vectorielle, « produit mixte » des vecteurs $f_1(t), \ldots, f_n(t)$. Il a pour éléments les fonctions dérivables $f_{jk}(t)$, donc est lui-même dérivable. En le dérivant d'après la règle 7.14*f* on trouve

$$\frac{d}{dt}[f_1(t), \ldots, f_n(t)] = [f_1'(t), f_2(t), \ldots, f_n(t)] + [f_1(t), f_2'(t), \ldots, f_n(t)] + \ldots + [f_1(t), \ldots, f_n'(t)] =$$
$$= [A(t) f_1(t), f_2(t), \ldots, f_n(t)] + [f_1(t), A(t) f_2(t), \ldots, f_n(t)] + \ldots$$
$$\ldots + [f_1(t), f_2(t), \ldots, A(t) f_n(t)].$$

Dans le k-ième terme de la somme trouvée, la seule composante du vecteur $A(t) f_k(t)$ qui importe est celle suivant le vecteur $f_k(t)$, car chaque autre composante amène à un déterminant ayant deux colonnes identiques, donc nul. La composante mentionnée vaut $a_{kk}(t) f_k(t)$. Définitivement, nous avons

$$\frac{d}{dt}[f_1(t), \ldots, f_n(t)] = (a_{11}(t) + \ldots + a_{nn}(t))[f_1(t), \ldots, f_n(t)]. \quad (2)$$

La quantité $\operatorname{tr} A(t) = a_{11}(t) + \ldots + a_{nn}(t)$ est la trace de la matrice $A(t)$. En intégrant l'équation (2) nous trouvons

$$[f_1(t), \ldots, f_n(t)] = [f_1(t_0), \ldots, f_n(t_0)]\, e^{\int_{t_0}^{t} \operatorname{tr} A(\tau)\, d\tau} \quad (3)$$

Rappelons que la quantité $\operatorname{tr} A(t)$ ne dépend pas du choix de la base $f_1, \ldots, f_n$ *).

13.74. Equation du n-ième ordre. Nous avons vu dans 13.51 que l'équation

$$y^{(n)}(t) = a_n(t) y^{(n-1)}(t) + \ldots + a_1(t) y'(t) + b(t) \quad (1)$$

est équivalente au système

$$\left.\begin{aligned} u_1'(t) &= u_2(t), \\ u_2'(t) &= u_3(t), \\ &\cdots\cdots\cdots \\ u_n'(t) &= a_1(t) u_1(t) + a_2(t) u_2(t) + \ldots + a_n(t) u_n(t) + b(t), \end{aligned}\right\} \quad (2)$$

où

$$u_1(t) = y(t), \quad u_2(t) = y'(t), \ldots, u_n(t) = y^{(n-1)}(t).$$

Ainsi, un vecteur solution $u_1(t), \ldots, u_n(t)$ correspond à une collection $y(t), y'(t), \ldots, y^{(n-1)}(t)$. Conformément à 13.72, toute solution $w(t)$ de l'équation homogène

$$w^{(n)}(t) = a_n(t) w^{(n-1)}(t) + \ldots + a_1(t) w_1(t) \quad (3)$$

*) Cf., par exemple, [14; 5.53].

peut être mise, d'une façon univoque, sous la forme

$$w(t) = C_1 w_1(t) + \ldots + C_n w_n(t),$$

où $w_1(t), \ldots, w_n(t)$ est la solution correspondant à une matrice non dégénérée des données initiales

$$\begin{Vmatrix} w_1(t_0) & \ldots & w_n(t_0) \\ \cdot & \cdot & \cdot \\ w_1^{(n-1)}(t_0) & \ldots & w_n^{(n-1)}(t_0) \end{Vmatrix} = W[w_1(t_0), \ldots, w_n(t_0)].$$

La matrice des solutions (matrice de Wronski)

$$\begin{Vmatrix} w_1(t) & \ldots & w_n(t) \\ \cdot & \cdot & \cdot \\ w_1^{(n-1)}(t) & \ldots & w_n^{(n-1)}(t) \end{Vmatrix} = W[w_1(t), \ldots, w_n(t)]$$

reste alors non dégénérée pour tout $t \in [a, b]$. Son déterminant, d'après la formule 13.73 (3) et compte tenu de la forme particulière de la matrice du système (2), vaut

$$\det W[w_1(t), \ldots, w_n(t)] = \det W[w_1(t_0), \ldots, w_n(t_0)]\, e^{\int_{t_0}^{t} a_n(\tau)\,d\tau}$$

§ 13.8. Résolution d'une équation linéaire non homogène

13.81. L'équation

$$u'(t) = A(t)\,u(t) + b(t) \tag{1}$$

avec $b(t) \not\equiv 0$ s'appelle *équation linéaire non homogène*. La différence $v_1(t) - v_2(t)$ de deux solutions $v_1(t)$ et $v_2(t)$ d'une équation non homogène est, évidemment, une solution de l'équation homogène correspondante. Par conséquent, étant donnée une solution particulière $v_1(t)$ de l'équation non homogène (celle, par exemple, qui vaut 0 pour $t = t_0$) et connaissant l'opérateur $\Omega_{t_0}^t$, on peut obtenir n'importe quelle solution de l'équation non homogène (celle, par exemple, qui vaut v_0 pour $t = t_0$) à l'aide de la formule

$$v(t) = v_1(t) + \Omega_{t_0}^t v_0.$$

Etant donné l'opérateur $\Omega_{t_0}^t$, la solution $v_1(t)$ peut être construite par la méthode de variation de la constante. Notamment, cherchons $v_1(t)$ sous la forme

$$v_1(t) = \Omega_{t_0}^t C(t), \tag{2}$$

où $C(t)$ est un vecteur variable inconnu (s'il ne dépendait pas de t, on aurait une solution de l'équation homogène, ce qui explique l'appellation de la méthode). Pour obtenir $v_1(t_0) = 0$, nous allons construire le vecteur $C(t)$ de façon que $C(t_0) = 0$.

En vertu de 13.38*e* et du lemme 13.13, nous avons

$$v_1'(t) = (\Omega_{t_0}^t)' C(t) + \Omega_{t_0}^t C'(t) = A(t)\,\Omega_{t_0}^t C(t) + \Omega_{t_0}^t C'(t). \quad (3)$$

D'autre part,

$$A(t)\,v_1(t) + b(t) = A(t)\,\Omega_{t_0}^t C(t) + b(t). \quad (4)$$

En égalant les seconds membres de (3) et (4), nous trouvons

$$\Omega_{t_0}^t C'(t) = b(t). \quad (5)$$

En appliquant aux deux membres de l'égalité l'opérateur $\Omega_t^{t_0}$ inverse de $\Omega_{t_0}^t$ (13.38*d*), nous obtenons

$$C'(t) = \Omega_t^{t_0} b(t),$$

d'où

$$C(t) = \int_{t_0}^{t} \Omega_\tau^{t_0} b(\tau)\,d\tau,$$

en choisissant le vecteur qui s'annule pour $t = t_0$.

Définitivement, nous aboutissons à la formule

$$v(t) = \Omega_{t_0}^t \int_{t_0}^{t} \Omega_\tau^{t_0} b(\tau)\,d\tau + \Omega_{t_0}^t v_0 = \Omega_{t_0}^t v_0 + \int_{t_0}^{t} \Omega_\tau^t b(\tau)\,d\tau. \quad (6)$$

Sa validité peut à présent être prouvée directement, en se servant de la règle de dérivation 9.86*b* (qui s'étend facilement aux fonctions vectorielles).

13.82. Si $B = R_1$ et $A(t)$ est une fonction numérique, nous avons (13.12) :

$$\Omega_{t_0}^t = e^{\int_{t_0}^{t} A(\tau)d\tau};$$

par conséquent

$$v(t) = e^{\int_{t_0}^{t} A(\tau)d\tau} v_0 + \int_{t_0}^{t} e^{\int_{\tau}^{t} A(\theta)d\theta} b(\tau)\,d\tau. \quad (1)$$

En particulier, pour $A(t)$ constante, $A(t) \equiv A$, on a la formule

$$v(t) = e^{(t-t_0)A} v_0 + \int_{t_0}^{t} e^{(t-\tau)A} b(\tau)\,d\tau \quad (2)$$

13.83. Les formules 13.81(6) et 13.82(1) sont également valables dans le cas général où $u(t)$ est une fonction vectorielle à valeurs

dans un espace normé B et A (t) un opérateur linéaire dans B. Les symboles figurant dans ces formules sont à considérer au sens de 13.13 et de 13.19, pour respectivement A $(t) = A$ constant et A (t) variable.

En particulier, si l'opérateur A $(t) = A$ ne dépend pas de t et si la fonction $b(t)$ a la forme particulière

$$b(t) = \sum_{k=1}^{m} P_k(t)\, e^{Q_k t} b_k,$$

où $P_k(t)$ est un polynôme, Q_k un opérateur constant qui commute avec l'opérateur A, b_k des vecteurs fixes de l'espace B, alors l'intégrale 13.82(2) peut être calculée explicitement. Dans le présent cas, la structure du résultat peut être prévue à des coefficients indéterminés près ; c'est pourquoi on peut chercher la solution par la méthode des coefficients indéterminés.

13.84. Considérons le cas d'espace n-dimensionnel $B = R_n$. Ici l'opérateur résolvant $\Omega_{t_0}^t$ est donné par la matrice de Wronski 13.72 :

$$\Omega_{t_0}^t = \begin{Vmatrix} f_{11}(t) & f_{12}(t) \dots f_{1n}(t) \\ f_{21}(t) & f_{22}(t) \dots f_{2n}(t) \\ \dots\dots\dots\dots\dots \\ f_{n1}(t) & f_{n2}(t) \dots f_{nn}(t) \end{Vmatrix}$$

On cherche la solution de la forme 13.81(2) ; dans le cas considéré on peut écrire

$$v_{1j}(t) = \sum_{k=1}^{n} f_{jk}(t)\, C_k(t),$$

où $C_1(t), \dots, C_n(t)$ sont les fonctions à déterminer. L'équation 13.81(5) prend la forme du système

$$\sum_{k=1}^{n} f_{jk}(t)\, C_k'(t) = b_k(t).$$

En le résolvant par rapport aux $C_k'(t)$ et en intégrant ensuite de t_0 à t on aura les quantités cherchées $C_k(t)$.

13.85. Considérons l'équation du n-ième ordre

$$y^{(n)}(t) = a_n(t)\, y^{(n-1)}(t) + \dots + a_1(t)\, y(t) + b(t).$$

L'opérateur résolvant du système équivalent

$$\begin{aligned} u_1'(t) &= u_2(t), \\ u_2'(t) &= u_3(t), \\ &\dots\dots\dots\dots\dots\dots \\ u_n'(t) &= a_1(t)\, u_1(t) + \dots + a_n(t)\, u_n(t), \end{aligned}$$

où $u_1(t) = y(t)$, $u_2(t) = y'(t)$, ..., $u_n(t) = y^{(n-1)}(t)$, a la forme (13.74)

$$\Omega_{t_0}^t = \begin{Vmatrix} w_1(t) & \dots & w_n(t) \\ w_1'(t) & \dots & w_n'(t) \\ \dots & \dots & \dots \\ w_1^{(n-1)}(t) & \dots & w_n^{(n-1)}(t) \end{Vmatrix},$$

$w_1(t)$, ..., $w_n(t)$ étant des solutions associées à une matrice non dégénérée des données initiales. La solution est cherchée sous la forme 13.81(2), c'est-à-dire que l'on a (pour la première ligne):

$$v(t) = \sum_{k=1}^{n} w_k(t)\, C_k(t).$$

Les équations 13.81(5) ont, dans le présent cas, la forme suivante:

$$\sum_{k=1}^{n} w_k(t)\, C_k'(t) = 0,$$

$$\sum_{k=1}^{n} w_k'(t)\, C_k'(t) = 0,$$

$$\dots\dots\dots\dots$$

$$\sum_{k=1}^{n} w_k^{(n-1)}(t)\, C_k'(t) = b(t).$$

En les résolvant par rapport aux $C_k'(t)$ et en intégrant de t_0 à t nous trouvons les fonctions $C_k(t)$, donc la solution cherchée $v(t)$.

Exercices

1. Nous avons deux solutions différentes $y \equiv 0$ et $y = x^3$ d'une même équation différentielle $\frac{dy}{dx} = 3y^{2/3}$, avec la même condition initiale $y(0) = 0$. Est-ce que ce fait ne contredit pas le théorème d'unicité 13.35?

2. Un point pesant P glisse sans frottement le long d'une courbe. Quelle doit être la courbe pour que la projection du point sur (a) la droite horizontale, (b) la droite verticale se déplace d'une façon uniforme?

3. On connaît une solution particulière $u_1(t)$ d'une équation linéaire du second ordre $u''(t) + a(t)\, u'(t) + b(t)\, u(t) = 0$. Comment trouver une autre solution; linéairement indépendante de la première?

4. Conformément à 13.17, une équation linéaire du n-ième ordre à coefficients constants est équivalente à un système du 1er ordre dont la matrice a un polynôme minimal de degré n. Montrer que tout système de n équations du 1er ordre possédant cette propriété est équivalent à une équation du n-ième ordre.

5. Soient $u'(t) = A(t)\, u(t)$ une équation vectorielle ($u(t) \in R_n$) et $A(t)$ un opérateur périodique de période T (i.e. $A(t+T) = A(t)$). Montrer que $\Omega_0^{t+T} = C\Omega_0^t$, où C est un opérateur constant.

6. On donne $n \leqslant N$ fonctions vectorielles linéairement indépendantes $u_1(t)$, ..., $u_n(t)$ à valeurs dans l'espace R_N et dérivables pour tout $t \in [a, b]$.

Montrer qu'il existe une équation $u'(t) = A(t)\,u(t)$ avec un opérateur continu $A(t)$ dans R_N qui a ces fonctions vectorielles pour solutions.

7. On a n fonctions scalaires linéairement indépendantes $y_1(t), \ldots, y_n(t)$ qui sont n fois dérivables. Est-il possible que leur déterminant de Wronski soit identiquement nul?

8. On a n fonctions scalaires linéairement indépendantes $y_1(t), \ldots, y_n(t)$ qui sont n fois dérivables et dont le wronskien est non nul. Construire une équation du n-ième ordre avec les solutions $y_1(t), \ldots, y_n(t)$.

9. Si les fonctions $A(t)$ et $b(t)$ ont les dérivées continues y compris la m-ième, alors une solution de l'équation linéaire

$$u'(t) = A(t)\,u(t) + b(t)$$

a les dérivées continues y compris la $(m+1)$-ième.

10. Si $y(0) = 0$ et l'inégalité $y'(t) - ky(t) \leqslant \varphi(t)$ a lieu pour $0 \leqslant t \leqslant T$, alors l'inégalité

$$y(t) \leqslant \int_0^t e^{-k(t-s)}\varphi(s)\,ds$$

a lieu elle aussi pour $0 \leqslant t \leqslant T$.

11. (Suite.) Si l'inégalité

$$w(t) \leqslant \varphi(t) + k\int_0^t w(s)\,ds$$

a lieu pour $0 \leqslant t \leqslant T$, alors il en est de même de l'inégalité

$$w(t) \leqslant \varphi(t) + k\int_0^t \varphi(s)\,e^{k(t-s)}\,ds.$$

12. (Suite.) Nous dirons qu'une fonction $y(t) \in X$ sur $[0, T]$ est une ε-presque-solution de l'équation $u'(t) = f(t, u)$ si

$$\| y'(t) - f(t, y(t)) \| \leqslant \varepsilon$$

sur $[0, T]$. Montrer que, pour toute solution $u(t)$ et pour toute ε-presque-solution $y(t)$, on a l'inégalité

$$\| u(t) - y(t) \| \leqslant \| u(0) + v(0) \|\, e^{kt} + \frac{\varepsilon}{k}\,[e^{kt} - 1],$$

où k est la constante dans la condition de Lipschitz pour la fonction $f(t, u)$.

13. (Suite.) Considérons l'équation

$$u'(t) = A(t)\,u(t), \quad u(0) = u_0, \quad 0 \leqslant t \leqslant T, \tag{1}$$

avec un opérateur continu donné $A(t)$. Soit

$$\Pi = \{0 = t_0 < t_1 < \ldots < t_n = T\}.$$

Définissons une fonction vectorielle continue $y_\Pi(t)$ comme suit:

$$y_\Pi(0) = y_0;$$

$$y_\Pi(t_{k+1}) = y_\Pi(t_k) + A(t_k)\,y_\Pi(t_k)\,\Delta t_k \quad (k = 0, 1, \ldots, n-1);$$

$y_\Pi(t)$ est du premier degré pour $t_k \leqslant t \leqslant t_{k+1}$ $(k = 0, 1, \ldots, n-1)$.

Pour un $\varepsilon > 0$ donné, construire une partition Π telle que la fonction $y_\Pi(t)$ devienne ε-presque-solution de l'équation (1).

14. (Suite.) Montrer qu'une solution de l'équation (1) (exercice 13) est donnée par l'expression

$$u(T) = \lim_{d(\Pi)\to 0} y_\Pi(T) = \lim_{d(\Pi)\to 0} \left\{ \prod_{k=n-1}^{0} (E + A(t_k)\,\Delta t_k) \right\} u_0$$

(l'ordre des facteurs est important!).

15. Dans l'hypothèse de l'exercice 13, définir une fonction vectorielle continue $z_\Pi(t)$ d'après la règle

$$z_\Pi(0) = y_0 ;$$

$$z_\Pi(t_{k+1}) = \left\{ \prod_{j=k}^{0} e^{A(t_j)\Delta t_j} \right\} y_0, \qquad (k = 0, 1, \ldots, n-1) ;$$

$$z_\Pi(t) = e^{A(t_k)(t-t_k)} z_\Pi(t_k), \qquad (t_k \leqslant t \leqslant t_{k+1}).$$

Pour un $\varepsilon > 0$ donné, construire une partition Π telle que la fonction $z_\Pi(t)$ devienne ε-presque-solution de l'équation (1).

16. Montrer qu'une solution de l'équation (1) (exercice 13) est donnée par l'expression

$$u(T) = \lim_{d(\Pi)\to 0} z_\Pi(T) = \lim_{d(\Pi)\to 0} \left\{ \prod_{j=n-1}^{0} e^{A(t_j)\Delta t_j} \right\} u_0$$

(l'ordre des facteurs est essentiel!).

Historique

Certaines équations différentielles apparaissent en mathématiques depuis la découverte du calcul différentiel et intégral, c'est-à-dire depuis les travaux de Newton et Leibniz. Leibniz intègre, en 1693, une équation linéaire homogène du premier ordre. La solution d'une équation linéaire, homogène ou non, du n-ième ordre à coefficients constants est trouvée par Euler (1739). La méthode de variation de la constante est élaborée par Lagrange (1775) ; d'ailleurs, Euler résout de divers problèmes par cette méthode à partir de 1739.

Durant le XVIIIe siècle, la théorie des équations différentielles détermine des progrès décisifs dans la mécanique ordinaire et céleste, la théorie des marées, la météorologie et d'autres domaines de la physique.

Les succès de la théorie des équations différentielles conditionnent la conclusion philosophique sur son caractère universel que l'on appelle « principe du déterminisme mécanique » et qui est exprimée dans l'épigraphe du présent chapitre. En mettant en relief le triomphe final de la raison, cette conclusion joue, à l'époque, un grand rôle dans la libération de la science des influences théologiques et scolastiques. Pourtant, les succès de la physique du XXe siècle montrent l'étroitesse du déterminisme mécanique qui, tout en conservant sa valeur pour les problèmes mécaniques, cède sa place dans les hauts domaines de la physique au déterminisme statistique.

Wronski, mathématicien et philosophe, introduit son déterminant des dérivées en 1812.

La position du problème général d'existence et d'unicité de la solution d'une équation différentielle est l'œuvre du XIXe siècle. La première démonstration de l'existence de la solution est donnée par Cauchy (1844) ; ensuite Lipschitz la simplifie essentiellement et formule la condition qui porte son nom. La méthode des approximations successives est proposée par Picard (1890) ; cette méthode est mise sous la forme abstraite pour un espace métrique, avec l'utilisation explicite d'un opérateur contractant, par Banach en 1922.

Développements orthogonaux

Le théorème de Fourier représente non seulement l'un des plus beaux résultats de l'analyse moderne, mais aussi un instrument indispensable pour l'étude de presque tous les problèmes principaux de la physique moderne.

W. Thomson et *P. G. Tait*,
Philosophie naturelle (1867)

§ 14.1. Développements orthogonaux dans un espace hilbertien

14.11. Position du problème. Dans le § 12.5, nous nous sommes occupés des approximations uniformes dans l'espace $C(Q)$ des fonctions continues, c'est-à-dire des approximations par rapport à la norme

$$\|f\|_C = \max_Q |f(x)|. \tag{1}$$

Dans plusieurs problèmes de l'analyse, il importe de considérer les approximations au sens de la moyenne intégrale, i.e. par rapport à la norme

$$\|f\|_p = \sqrt[p]{\int_Q |f(x)|^p \, dx}, \tag{2}$$

et, en premier lieu, les approximations au sens de la moyenne quadratique correspondant à la norme

$$\|f\|_2 = \sqrt{\int_Q |f(x)|^2 \, dx}. \tag{3}$$

Dans le présent chapitre, nous sous-entendons par Q un intervalle fermé de l'axe réel.

Quelles sont les conditions à imposer à un système (linéaire) donné $B(Q)$ de fonctions $\varphi(x)$ pour que, quelle que soit une fonction $f(x)$ (continue ou au moins continue par morceaux, ce qui garantit l'existence des intégrales nécessaires), on puisse indiquer une suite $\varphi_1(x)$, $\varphi_2(x)$, ... de fonctions de $B(Q)$ telle que

$$\|f - \varphi_n\|_2 = \sqrt{\int_Q |f(x) - \varphi_n(x)|^2 \, dx} \to 0 \tag{4}$$

lorsque $n \to \infty$?

Si une suite $\varphi_n(x)$ converge uniformément vers $f(x)$, alors la relation (4) est évidemment valable. Mais la convergence uniforme ne

découle pas de la relation (4), ni même la convergence simple. On peut donc dire que le problème d'approximations au sens de la moyenne quadratique est « plus facile » que celui d'approximations uniformes. De plus, la construction d'approximations du premier type peut être mise sous une forme géométrique bien claire du fait que l'espace correspondant de fonctions avec la norme (3) est un espace hilbertien où l'on peut non seulement mesurer les longueurs des vecteurs comme dans un espace normé, mais aussi utiliser la propriété d'orthogonalité.

14.12. Approximations dans un espace hilbertien. Soit H un espace hilbertien réel ou complexe. Soit $B \subset H$ le sous-espace n-dimensionnel engendré par un système orthonormé $e_1, \ldots, e_n$ (de sorte que $(e_j, e_k) = 1$ pour $j = k$ et $(e_j, e_k) = 0$ pour $j \neq k$). Posons le problème suivant: pour un vecteur $f \in H$ donné, trouver un vecteur $y = \sum_{k=1}^{n} c_k e_k \in B$ tel que la quantité $\| f - y \|$ soit la plus petite possible. En résolvant ce problème nous supposons l'espace H complexe; dans le cas réel, on n'aura qu'à simplifier un peu les notations. Pour tout $x = \sum_{k=1}^{n} \xi_k e_k$, nous avons

$$\| f - x \|^2 = \left(f - \sum_{k=1}^{n} \xi_k e_k,\ f - \sum_{k=1}^{n} \xi_k e_k\right) =$$

$$= (f, f) - \sum_{k=1}^{n} \xi_k \overline{(f, e_k)} - \sum_{k=1}^{n} \bar{\xi}_k (f, e_k) + \sum_{k=1}^{n} |\xi_k|^2 =$$

$$= (f, f) + \sum_{k=1}^{n} \left[|(f, e_k)|^2 - \xi_k \overline{(f, e_k)} - \bar{\xi}_k (f, e_k) + |\xi_k|^2 \right] -$$

$$- \sum_{k=1}^{n} |(f, e_k)|^2 =$$

$$= (f, f) + \sum_{k=1}^{n} |(f, e_k) - \xi_k| \overline{|(f, e_k) - \xi_k|} - \sum_{k=1}^{n} |(f, e_k)|^2 =$$

$$= (f, f) + \sum_{k=1}^{n} |(f, e_k) - \xi_k|^2 - \sum_{k=1}^{n} |(f, e_k)|^2. \qquad (1)$$

Il est évident que l'expression obtenue est minimale si $\xi_k = (f, e_k)$ $(k = 1, \ldots, n)$. Le vecteur $y = \sum_{k=1}^{n} (f, e_k)\, e_k$ s'appelle *projection du vecteur f sur le sous-espace B*, le vecteur $f - y = h$ étant la *perpendiculaire* abaissée de l'extrémité du vecteur f sur le sous-espace B; dans le cas réel, ces définitions s'accordent avec les

notions géométriques correspondantes (fig. 14.1). On voit de (1) que le carré de la longueur du vecteur h est

$$\| h \|^2 = (f, f) - \sum_{k=1}^{n} | (f, e_k) |^2.$$

Il en résulte, en particulier, l'*inégalité de Bessel*

$$\sum_{k=1}^{n} | (f, e_k) |^2 \leqslant \| f \|^2 \tag{2}$$

qui est valable pour tout vecteur $f \in H$ et tout système orthonormé $e_1, \ldots, e_n$.

Définitivement, la meilleure *approximation « hilbertienne »* du vecteur f par les vecteurs du sous-espace B est réalisée lorsque l'on choisit pour vecteur approximant y la projection du vecteur f sur le sous-espace B.

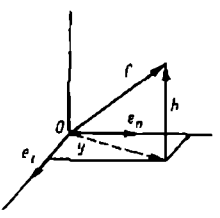


Fig. 14.1.

14.13. Soit maintenant un système orthonormé infini $e_1, e_2, \ldots, e_n, \ldots$ On peut effectuer la construction décrite plus haut pour toute famille finie $e_1, \ldots, e_n$ en obtenant la meilleure approximation hilbertienne correspondante $y_n = \sum_{k=1}^{n} (f, e_k) e_k$. Notons que les coefficients (f, e_k) de la meilleure approximation ne dépendent pas du numéro $n \geqslant k$. L'écart h_n vaut, comme nous l'avons vu,

$$\| h_n \| = \sqrt{(f, f) - \sum_{k=1}^{n} | (f, e_k) |^2}. \tag{1}$$

La question se pose : est-il possible de rendre la quantité $\| h_n \|$ aussi petite que l'on veut en choisissant n suffisamment grand ? Cela ne peut avoir lieu dans le cas général : par exemple, si un système $e_1, e_2, \ldots$ « n'est pas complet », i.e. s'il existe un vecteur f non nul et orthogonal à tous les vecteurs $e_1, e_2, \ldots$, alors tous les nombres (f, e_k) sont nuls et tous les nombres $\| h_n \|$ valent $\| f \|$.

14.14.a. Tout de même, dans certaines conditions on peut affirmer que les nombres $\| h_n \|$ tendent vers zéro lorsque $n \to \infty$. A savoir, ceci aura lieu si, en partant de quelques considérations supplémentaires (par exemple, des théorèmes du type Stone (12.52) pour certains espaces fonctionnels), on sait qu'il est possible de choisir, parmi les combinaisons linéaires des vecteurs $e_1, e_2, \ldots$, une suite convergente (par rapport à la norme hilbertienne) vers le vecteur f. En effet, si, pour une combinaison linéaire $\sum_{k=1}^{n} \xi_k e_k$,

on a l'inégalité $\| f - \sum_{k=1}^{n} \xi_k e_k \| \leqslant \varepsilon$, alors, pour la meilleure approximation hilbertienne $\sum_{k=1}^{n} (f, e_k) e_k$, on a à fortiori :

$$\| h_n \| = \sqrt{\| f \|^2 - \sum_{k=1}^{n} | (f, e_k) |^2} = \| f - \sum_{k=1}^{n} (f, e_k) e_k \| \leqslant \varepsilon. \tag{1}$$

b. Dans l'hypothèse a, en passant à la limite pour $\varepsilon \to 0$, on obtient l'égalité

$$f = \lim_{n \to \infty} \sum_{k=1}^{n} (f, e_k) e_k = \sum_{k=1}^{\infty} (f, e_k) e_k. \tag{2}$$

La série dans le dernier membre s'appelle *série de Fourier* du vecteur f suivant le système orthonormé $e_1, e_2, \ldots$ Les nombres (f, e_k) sont appelés *coefficients de Fourier* du vecteur f suivant le système e_k. Ces termes restent valables indépendamment de la convergence de la série ; en cas de divergence, nous considérons la série de Fourier pour le moment formellement.

c. En cas de convergence de la série (2) vers le vecteur f, nous avons en passant à la limite dans (1) :

$$\| f \|^2 = \sum_{k=1}^{\infty} | (f, e_k) |^2. \tag{3}$$

Cette égalité qui représente un analogue du théorème de Pythagore pour le cas de dimension infinie s'appelle *égalité de Parseval*. Si l'on ne peut rien dire sur la convergence de la série de Fourier vers le vecteur f, on a toujours l'inégalité

$$\sum_{k=1}^{\infty} | (f, e_k) |^2 \leqslant \| f \|^2 \tag{4}$$

qui se déduit par un passage à la limite dans l'inégalité de Bessel 14.12(2) et s'appelle elle aussi inégalité de Bessel.

d. Remarquons enfin que si la série de Fourier d'un vecteur f converge vers f,

$$f = \sum_{k=1}^{\infty} (f, e_k) e_k,$$

alors il en sera de même après n'importe quelle permutation des termes mettant le terme de numéro n_k à la place k ($k = 1, 2, \ldots$) :

$$f = \sum_{k=1}^{\infty} (f, e_{n_k}) e_{n_k}.$$

En effet, d'apres (1) on a

$$\Big\| f - \sum_{k=1}^{N} (f,\ e_{n_k})\, e_{n_k} \Big\|^2 = \| f \|^2 - \sum_{k=1}^{N} |(f,\ e_{n_k})|^2.$$

La dernière quantité tend vers zéro lorsque $N \to \infty$ en vertu de (3) et de la légitimité de permutations dans une série convergente à termes positifs (6.36).

14.15. La série de Fourier apparaît nécessairement dans les problèmes d'approximations puisque l'on a la propriété suivante:

L e m m e. *Supposons que, pour un vecteur* $f \in \mathrm{H}$ *et un système orthonormé* $\{e_k\}$ *dans* H, *on ait un développement*

$$f = \sum_{1}^{\infty} c_k e_k, \tag{1}$$

au sens que

$$\lim_{p \to \infty} \Big\| f - \sum_{k=1}^{n_p} c_k e_k \Big\| = 0$$

pour au moins une suite $n_1 < n_2 < \ldots < n_p < \ldots$. *Alors tout coefficient* c_m *est égal au coefficient correspondant de Fourier* $(f,\ e_m)$ $(m = 1,\ 2,\ \ldots)$, *et la série* (1) *converge en norme au sens ordinaire.*

D é m o n s t r a t i o n. En multipliant (1) scalairement par e_m et en utilisant la continuité du produit scalaire (12.43*b*) et le fait que le système $\{e_k\}$ est orthonormé, on obtient

$$(f,\ e_m) = \Big(\lim_{p \to \infty} \sum_{1}^{n_p} c_k e_k,\ e_m \Big) = \lim_{p \to \infty} \Big(\sum_{1}^{n_p} c_k e_k,\ e_m \Big) = c_m,$$

ce qu'il nous fallait.

14.16. L e m m e. *Soient*

$$f = \sum_{1}^{\infty} a_k e_k, \quad g = \sum_{1}^{\infty} b_k e_k$$

deux développements (au même sens que dans 14.15), $\{e_k\}$ *étant un système orthonormé; la série* $\sum_{1}^{\infty} a_k \overline{b}_k$ *converge absolument, et nous avons*

$$\sum_{1}^{\infty} a_k \overline{b}_k = (f,\ g). \tag{1}$$

D é m o n s t r a t i o n. La convergence absolue de la série dans le premier membre de (1) découle de l'inégalité

$$|a_k b_k| \leqslant \frac{1}{2} (|a_k|^2 + |b_k|^2)$$

en vertu de 14.14 (4). Ensuite, si

$$f = \lim_{p\to\infty} \sum_{1}^{m_p} a_k e_k, \quad g = \lim_{p\to\infty} \sum_{1}^{n_p} b_k e_k$$

en norme de l'espace H et si $l_p = \min(m_p, n_p)$, alors, en raison de la continuité du produit scalaire, on a bien

$$(f, g) = \Big(\lim_{p\to\infty} \sum_{1}^{m_p} a_k e_k, \lim_{p\to\infty} \sum_{1}^{n_p} b_k e_k\Big) =$$

$$= \lim_{p\to\infty} \sum_{1}^{l_p} a_k \overline{b}_k = \sum_{1}^{\infty} a_k \overline{b}_k.$$

14.17. Ecrivons, pour les utiliser par la suite, la série de Fourier et l'égalité de Parseval pour un système orthogonal mais non normé de vecteurs $g_1, g_2, \ldots$. Si le vecteur g_n est non nul pour tout n, alors le système $e_n = g_n / \| g_n \|$ est déjà orthonormé. La série de Fourier d'un vecteur f suivant le système e_n peut être écrite comme suit

$$\sum_{k=1}^{\infty} (f, e_k)\, e_k = \sum_{k=1}^{\infty} \Big(f, \frac{g_k}{\| g_k \|}\Big) \frac{g_k}{\| g_k \|} =$$

$$= \sum_{k=1}^{\infty} \frac{(f, g_k)}{\| g_k \|^2} g_k = \sum_{k=1}^{\infty} \alpha_k g_k, \tag{1}$$

où

$$\alpha_k = \frac{(f, g_k)}{\| g_k \|^2}. \tag{2}$$

La série dans le dernier membre de (1) s'appelle *série de Fourier du vecteur f suivant le système g_n*, les nombres α_k sont *les coefficients de Fourier du vecteur f suivant le système g_n*. Si la série (1) converge vers f, alors on a

$$\| f \|^2 = \sum_{k=1}^{\infty} |(f, e_k)|^2 = \sum_{k=1}^{\infty} \Big| \Big(f, \frac{g_k}{\| g_k \|^2}\Big) \Big|^2 =$$

$$= \sum_{k=1}^{\infty} \frac{|(f, g_k)|^2}{\| g_k \|^2} = \sum_{k=1}^{\infty} \| g_k \|^2 \alpha_k^2,$$

ce qui représente *l'égalité de Parseval pour le système g_n*. Et l'égalité 14.16(1) devient

$$(f, g) = \sum_{n=1}^{\infty} \alpha_k \overline{\beta}_k \| g_k \|^2 \tag{3}$$

en supposant que

$$f=\sum_1^\infty \alpha_k g_k, \quad g=\sum_1^\infty \beta_k g_k.$$

§ 14.2. Séries de Fourier classiques

14.21. Dans l'espace hilbertien réel $H_R[-\pi, \pi]$ (12.48*d*) des fonctions $f(t)$ continues par morceaux sur l'intervalle $[-\pi, \pi]$ ou, ce qui revient au même, sur la circonférence $Q=\{(x, y): x=\cos t,\ y=\sin t\}$, considérons le système orthogonal infini

$$1,\ \cos t,\ \sin t,\ \cos 2t,\ \sin 2t,\ \ldots,\ \cos nt,\ \sin nt,\ \ldots \qquad (1)$$

L'orthogonalité de ce système est immédiate en calculant les intégrales des fonctions $\cos kt \cdot \cos mt$, $\cos kt \cdot \sin mt$, $\sin kt \cdot \sin mt$ sur l'intervalle $[-\pi, \pi]$.

Pour normer le système, remarquons que

$$\|1\|^2=\int_{-\pi}^{\pi} 1^2\,dt=2\pi$$

et, d'après 9.55*c*,

$$\|\cos kt\|^2=\int_{-\pi}^{\pi}\cos^2 kt\,dt=\pi,$$

$$\|\sin kt\|^2=\int_{-\pi}^{\pi}\sin^2 kt\,dt=\pi.$$

Par conséquent, étant donnée une fonction $f(t)\in H_R[-\pi, \pi]$, la série de Fourier 14.17(1) s'écrit de la façon suivante:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\tau)\,d\tau+\cos t\cdot\frac{1}{\pi}\int_{-\pi}^{\pi} f(\tau)\cos\tau\,d\tau+$$
$$+\sin t\cdot\frac{1}{\pi}\int_{-\pi}^{\pi} f(\tau)\sin\tau\,d\tau+\ldots=$$
$$=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nt+b_n\sin nt), \qquad (2)$$

où

$$\left.\begin{aligned} a_n&=\frac{1}{\pi}\int_{-\pi}^{\pi} f(\tau)\cos n\tau\,d\tau,\\ b_n&=\frac{1}{\pi}\int_{-\pi}^{\pi} f(\tau)\sin n\tau\,d\tau \quad (n=0,\ 1,\ 2,\ \ldots).\end{aligned}\right\} \qquad (3)$$

Les nombres a_n et b_n s'appellent *coefficients de Fourier de la fonction $f(t)$ suivant le système* (1) *de fonctions trigonométriques.* Le problème de convergence de la série de Fourier (2) sera discuté plus bas (14.24).

14.22. Passons maintenant à une série trigonométrique de Fourier sous forme complexe. Considérons l'espace hilbertien complexe $H_C[-\pi, \pi]$ des fonctions complexes continues par morceaux sur l'intervalle $-\pi \leqslant t \leqslant \pi$ (12.48*e*). Ici on a un système orthogonal infini des fonctions

$$e^{int} \quad (n = 0, \pm 1, \pm 2, \ldots). \tag{1}$$

L'orthogonalité découle de l'égalité évidente

$$\int_{-\pi}^{\pi} e^{int}\overline{e^{imt}}\,dt = \int_{-\pi}^{\pi} e^{i(n-m)t}\,dt = \frac{e^{i(n-m)t}}{i(n-m)}\Big|_{-\pi}^{\pi} = 0 \quad (n \neq m).$$

Calculons la norme de la fonction e^{int} :

$$\|e^{int}\| = \sqrt{\int_{-\pi}^{\pi} |e^{int}|^2\,dt} = \sqrt{\int_{-\pi}^{\pi} 1 \cdot dt} = \sqrt{2\pi}.$$

Pour une fonction donnée $f(t) \in H_C[-\pi, \pi]$, la série de Fourier 14.17(1) prend la forme de la série bilatérale (6.48)

$$\sum_{n=-\infty}^{\infty} c_n e^{int}, \tag{2}$$

où

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\tau)\, e^{-in\tau}\,d\tau \tag{3}$$

sont les coefficients de Fourier de la fonction $f(t)$ suivant le système (1).

14.23. En utilisant la formule d'Euler

$$e^{int} = \cos nt + i \sin nt$$

on peut mettre la série 14.22(2) sous la forme 14.21(2) (plus haut, nous n'avons écrit la série 14.21(2) que pour les fonctions réelles). Posons

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\tau)\,d\tau = \frac{a_0}{2},$$

$$c_n e^{int} + c_{-n} e^{-int} = (c_n + c_{-n}) \cos nt + i\,(c_n - c_{-n}) \sin nt =$$

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau) \cos n\tau\,d\tau \cdot \cos nt + \frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau) \sin n\tau\,d\tau \cdot \sin nt =$$

$$= a_n \cos nt + b_n \sin nt.$$

Pour tout n, la somme partielle $\sum_{0}^{n}$ de la série 14.21(2) et la somme partielle symétrique $\sum_{-n}^{n}$ de 14.22(2) se confondent. Par conséquent, la convergence de la série 14.21(2) est équivalente à la sommabilité symétrique de la série 14.22(2). Soulignons que, comme nous l'avons remarqué dès 6.49, la sommabilité symétrique de la série 14.22(2) peut avoir lieu pour tout $t \in R_1$ sans que la série soit convergente (au sens de l'existence de $\lim_{m,\, n \to \infty} \sum_{-m}^{n}$).

14.24. Théorème. *La série de Fourier* 14.22(2) *de toute fonction (complexe) $f(t)$ continue par morceaux sur le compact $Q = [-\pi, \pi] = \{x^2 + y^2 = 1\}$ converge vers $f(t)$ en norme de l'espace $H_C(Q)$ quel que soit l'ordre de ses termes.*

Démonstration. Compte tenu du résultat 14.14*a*, il suffit de montrer qu'il existe une suite $T_n(t)$ de polynômes trigonométriques qui converge vers $f(t)$ en norme de l'espace $H_C(Q)$.

Considérons les polynômes trigonométriques

$$T_n(t) = \int_{-\pi}^{\pi} D_n(\tau;\ t)\, f(\tau)\, d\tau,$$

où $D_n(\tau;\ t)$ est la suite en forme de delta de 12.57*a*. D'après le théorème 12.57*c*, les polynômes $T_n(t)$ convergent vers $f(t)$ uniformément sur tout ensemble fermé $E \subset Q$ de points de continuité de la fonction $f(t)$. Ils restent bornés en module sur tout le Q par le nombre $M = \sup |f(t)|$. Pour un $\varepsilon > 0$ donné, l'ensemble fini des points de discontinuité de la fonction $f(t)$ peut être recouvert par un ensemble ouvert U, réunion d'un nombre fini d'intervalles dont la somme des longueurs est inférieure à ε^2. La fonction $f(t)$ est continue sur l'ensemble fermé $Q - U$. Trouvons un numéro n de façon à avoir $|f(t) - T_n(t)| < \varepsilon$ sur $Q - U$. Alors

$$\|f(t) - T_n(t)\|^2 = \int_Q |f(t) - T_n(t)|^2\, dt =$$

$$= \int_U |f(t) - T_n(t)|^2\, dt + \int_{Q-U} |f(t) - T_n(t)|^2\, dt \leqslant$$

$$\leqslant 4M^2\varepsilon^2 + 4\pi^2\varepsilon^2 = 4\varepsilon^2 (M^2 + \pi^2),$$

d'où

$$\lim_{n \to \infty} T_n(t) = f(t)$$

dans $H_C(Q)$, ce qu'il nous fallait.

D'après ce qu'on a dit plus haut, le fait analogue a lieu pour toute fonction réelle continue par morceaux $f(t)$ et sa série de Fourier 14.21(2).

14.25. Du même coup, nous obtenons l'égalité de Parseval pour toute fonction $f(t)$ continue par morceaux : dans le cas réel,

$$\int_{-\pi}^{\pi} f^2(t)\,dt = \pi\frac{a_0^2}{2} + \pi\sum_{1}^{\infty}(a_k^2 + b_k^2), \tag{1}$$

d'après les formules 14.17 et 14.21 ; dans le cas complexe,

$$\int_{-\pi}^{\pi} |f(t)|^2\,dt = 2\pi\sum_{-\infty}^{\infty}|c_n|^2, \tag{2}$$

d'après les formules 14.17 et 14.22.

Etant données deux fonctions continues par morceaux $f(t)$ et $g(t)$, nous voyons de 14.17(3) que, dans le cas réel où

$$f(t) = \frac{a_0}{2} + \sum_{1}^{\infty}(a_n\cos nt + b_n\sin nt),$$

$$g(t) = \frac{c_0}{2} + \sum_{1}^{\infty}(c_n\cos nt + d_n\sin nt),$$

on a l'égalité

$$\int_{-\pi}^{\pi} f(t)\,g(t)\,dt = \pi\frac{a_0c_0}{2} + \pi\sum_{1}^{\infty}(a_nc_n + b_nd_n), \tag{3}$$

et dans le cas complexe où

$$f(t) = \sum_{-\infty}^{\infty} a_ne^{int}, \quad g(t) = \sum_{-\infty}^{\infty} b_ne^{int},$$

on a

$$\int_{-\pi}^{\pi} f(t)\,\overline{g(t)}\,dt = 2\pi\sum_{-\infty}^{\infty} a_n\overline{b_n}. \tag{4}$$

14.26. Le théorème 14.24 et les égalités 14.25(1)-(2) impliquent, en particulier, les résultats suivants :

a. Conséquence. *Les coefficients de Fourier* a_n, b_n *d'une fonction continue par morceaux* $f(t)$ *tendent vers* 0 *pour* $n \to \infty$, *les coefficients* c_n *tendent vers* 0 *pour* $n \to \pm\infty$.

b. Conséquence. *Si tous les coefficients de Fourier* 14.21(3) *(ou* 14.22(3)*) d'une fonction continue par morceaux* $f(t)$ *sont nuls, alors* $f(t)$ *est nulle partout, à l'exception possible d'un nombre fini de points* (9.16e).

c. Conséquence. *Deux fonctions continues par morceaux $f(t)$ et $g(t)$ dont les coefficients de Fourier* 14.21(3) *ou* 14.22(3) *sont respectivement égaux se confondent partout, à l'exception possible d'un nombre fini de points.*

d. Conséquence. *Le système orthogonal* 14.21(1) *des fonctions trigonométriques est complet : toute fonction continue par morceaux $f(t)$ qui est orthogonale à toutes les fonctions du système* 14.21(1) *représente le zéro de l'espace $H_R(Q)$, i.e. est nulle partout, à l'exception possible d'un nombre fini de points. Il en est de même du système des fonctions e^{int}* (14.22(1)) *dans l'espace $H_C(Q)$.*

14.27.a. Lemme. *Si une série trigonométrique $\sum_{-\infty}^{\infty} c_k e^{ikt}$ converge vers une fonction $s(t)$ uniformément sur $[-\pi, \pi]$, au sens que, pour certaines suites $m_p \to \infty$, $n_p \to \infty$, on a*

$$s(t) = \lim_{p\to\infty} \sum_{-m_p}^{n_p} c_k e^{ikt}$$

uniformément sur $[-\pi, \pi]$, alors les nombres c_k coïncident avec les coefficients de Fourier de la fonction $s(t)$.

Le lemme est immédiat parce que la fonction $s(t)$ est évidemment continue et la convergence uniforme implique la convergence en norme de $H_C(Q)$, ce qui permet d'appliquer 14.15.

b. Conséquence. *Si c_n ($n = 0, \pm 1, \ldots$) sont les coefficients de Fourier d'une fonction continue par morceaux $f(t)$ et si la série $\sum_{-\infty}^{\infty} c_n e^{int}$ converge uniformément vers une fonction $s(t)$ au même sens que dans a, alors $f(t) \equiv s(t)$ partout, à l'exception possible d'un nombre fini de points.*

En effet, la fonction $s(t)$ est évidemment continue et les nombres c_n sont, en vertu de a, ses coefficients de Fourier ; or les mêmes nombres sont, par hypothèse, les coefficients de Fourier de la fonction $f(t)$. L'application de la conséquence 14.26c conduit au résultat cherché.

§ 14.3. Convergence d'une série de Fourier en un point et sur un ensemble

14.31. Convergence d'une série de Fourier en un point. Nous avons vu dans 14.24 que la série de Fourier permet d'obtenir une approximation illimitée d'une fonction continue par morceaux donnée $f(t)$ au sens de la moyenne quadratique. Nous voulons savoir si la série de Fourier permet d'obtenir une approximation illimitée de la valeur de la fonction $f(t)$ en un point donné $t = t_0$, en d'autres termes, si la série de Fourier converge, pour $t = t_0$, vers le nombre $f(t_0)$ au sens ordinaire.

A cette fin, considérons la série de Fourier sous sa forme complexe 14.22(2). Soit

$$s_{m,\,n} = \sum_{k=-m}^{n} c_k e^{ikt}$$

une somme partialle de la série de Fourier d'une fonction $f(t)$, pour certains $m > 0$, $n > 0$. Nous avons

$$s_{m,\,n}(t) = \sum_{-m}^{n} c_k e^{ikt} = \frac{1}{2\pi} \sum_{-m}^{n} \int_{-\pi}^{\pi} f(\tau)\, e^{-ik\tau}\, d\tau \cdot e^{ikt} =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\tau) \sum_{-m}^{n} e^{ik(t-\tau)}\, d\tau.$$

En sommant la progression géométrique, nous obtenons

$$\sum_{-m}^{n} e^{ik\theta} = \frac{e^{-im\theta} - e^{i(n+1)\theta}}{1 - e^{i\theta}} = \frac{e^{i\left(n+\frac{1}{2}\right)\theta} - e^{-i\left(m+\frac{1}{2}\right)\theta}}{e^{i\frac{\theta}{2}} - e^{-i\frac{\theta}{2}}} =$$

$$= \frac{e^{i\left(n+\frac{1}{2}\right)\theta} - e^{-i\left(m+\frac{1}{2}\right)\theta}}{2i \sin\frac{\theta}{2}},$$

donc

$$\frac{1}{4\pi i} \int_{-\pi}^{\pi} f(\tau)\, \frac{e^{i\left(n+\frac{1}{2}\right)(t-\tau)} - e^{-i\left(m+\frac{1}{2}\right)(t-\tau)}}{\sin\frac{t-\tau}{2}}\, d\tau =$$

$$= \frac{1}{4\pi i} \int_{-\pi}^{\pi} f(t+h)\, \frac{e^{i\left(n+\frac{1}{2}\right)h} - e^{-i\left(m+\frac{1}{2}\right)h}}{\sin\frac{h}{2}}\, dh. \qquad (1)$$

Si l'on pose $f(t) \equiv 1$, on a évidemment $s_{m,\,n}(t) \equiv f(t)$ pour n'importe quels $m > 0$ et $n > 0$. Dans ce cas la formule (1) donne

$$\frac{1}{4\pi i} \int_{-\pi}^{\pi} \frac{e^{i\left(n+\frac{1}{2}\right)h} - e^{-i\left(m+\frac{1}{2}\right)h}}{\sin\frac{h}{2}}\, dh = 1.$$

A présent, la différence $s_{m,\,n}(t) - f(t)$ prend la forme

$$s_{m,\,n}(t) - f(t) =$$

$$= \frac{1}{4\pi i} \int_{-\pi}^{\pi} \{f(t+h) - f(t)\}\, \frac{e^{i\left(n+\frac{1}{2}\right)h} - e^{-i\left(m+\frac{1}{2}\right)h}}{\sin\frac{h}{2}}\, dh. \qquad (2)$$

Dans quelles conditions $s_{m,n}(t)$ tend vers $f(t)$ ou, ce qui revient au même, l'intégrale (2) tend vers zéro.

14.32. Lemme. *Si $\varphi(h)$ est une fonction à valeurs complexes, définie dans l'intervalle $a < x \leqslant b$, bornée et continue par morceaux sur tout intervalle $[a+\delta, b]$, $\delta > 0$, absolument intégrable au sens impropre sur $[a, b]$, alors les intégrales*

$$\int_a^b \varphi(h) \sin \nu h \, dh, \quad \int_a^b \varphi(h) \cos \nu h \, dh, \quad \int_a^b \varphi(h) e^{i\nu h} dh$$

tendent vers zéro lorsque $\nu \to \pm\infty$.

Démonstration. Vu les formules $\sin \nu h = \frac{1}{2i}(e^{i\nu h} - e^{-i\nu h})$, $\cos \nu h = \frac{1}{2}(e^{i\nu h} + e^{-i\nu h})$, il suffit de considérer la dernière intégrale. Envisageons d'abord le cas où la fonction $\varphi(h)$ est continue sur l'intervalle $[a, b]$. Pour ν fixe, il n'y a sur cet intervalle qu'un nombre fini de points de la progression arithmétique

$$a + k\frac{2\pi}{\nu}, \quad k = 0, 1, 2, \ldots;$$

désignons-les par $h_0 = a < h_1 < \ldots < h_m$. Supposons que la fonction $g(h)$ vaille $\varphi(h_0)$ dans l'intervalle $[h_0, h_1)$, $\varphi(h_1)$ dans $[h_1, h_2)$, etc., enfin $\varphi(h_m)$ dans l'intervalle $[h_m, b)$. Il est évident que

$$|g(h) - \varphi(h)| \leqslant \omega_\varphi\left(\frac{2\pi}{\nu}\right) \tag{1}$$

partout sur $[a, b]$, $\omega_\varphi(\delta)$ étant l'oscillation de la fonction $\varphi(x)$ sur $[a, b]$ (5.17*c*). Comme l'intervalle $[h_j, h_{j+1}]$ est une période de la fonction $e^{i\nu h}$, on a

$$\int_{h_j}^{h_{j+1}} g(h) e^{i\nu h} dh = \varphi(h_j) \int_{h_j}^{h_{j+1}} e^{i\nu h} dh = 0;$$

donc

$$\left| \int_a^b g(h) e^{i\nu h} dh \right| = \left| \int_{h_m}^b g(m) e^{i\nu h} dh \right| \leqslant M \cdot \frac{2\pi}{\nu}, \tag{2}$$

où $M = \max |\varphi(h)|$. Il résulte de (1) et (2) que

$$\left| \int_a^b \varphi(h) e^{i\nu h} dh \right| \leqslant \int_a^b |\varphi(h) - g(h)| \, dh + \left| \int_a^b g(h) e^{i\nu h} dh \right| \leqslant$$

$$\leqslant \omega_\varphi\left(\frac{2\pi}{\nu}\right)(b-a) + M\frac{2\pi}{\nu}. \tag{3}$$

Puisque, en vertu de la continuité de $\varphi(h)$, la quantité $\omega\left(\frac{2\pi}{\nu}\right)$ tend vers zéro lorsque $\nu \to \infty$, l'estimation (3) démontre le lemme dans le cas considéré.

Soit maintenant $\varphi(h)$ une fonction quelconque vérifiant l'hypothèse du lemme. Pour un $\varepsilon > 0$ donné, trouvons un $\delta > 0$ de façon à avoir

$$\int_a^{a+\delta} |\varphi(h)|\, dh < \frac{\varepsilon}{3}. \tag{4}$$

Sur l'intervalle $[a + \delta, b]$, la fonction $\varphi(h)$ est bornée en module par un nombre $M = M(\varepsilon)$ et continue par morceaux. L'intervalle $[a + \delta, b]$ se décompose en un nombre fini, soit $N = N(\varepsilon)$, d'intervalles $[a_1, b_1], \ldots, [a_N, b_N]$ sur chacun desquels la fonction $\varphi(h)$ est continue. En appliquant à chaque intervalle l'estimation (3) et compte tenu de (4), on trouve

$$\left|\int_a^b \varphi(h)\, e^{i\nu h}\, dh\right| \leqslant \int_a^{a+\delta} |\varphi(h)|\, dh + \sum_{k=1}^{N} \left|\int_{a_k}^{b_k} \varphi(h)\, e^{i\nu h}\, dh\right| \leqslant$$

$$\leqslant \frac{\varepsilon}{3} + \omega_\varphi\left(\frac{2\pi}{\nu}\right) \sum_{k=1}^{N} (b_k - a_k) + N(\varepsilon)\, M(\varepsilon)\, \frac{2\pi}{\nu} \leqslant$$

$$\leqslant \frac{\varepsilon}{3} + \omega_\varphi\left(\frac{2\pi}{\nu}\right)(b-a) + N(\varepsilon)\, M(\varepsilon)\, \frac{2\pi}{\nu}. \tag{5}$$

Il reste maintenant de trouver d'après ε donné un ν tel que les quantités $\omega_\varphi\left(\frac{2\pi}{\nu}\right)(b-a)$ et $N(\varepsilon)\, M(\varepsilon)\, \frac{2\pi}{\nu}$ deviennent inférieures à $\varepsilon/3$. Ceci fait, le lemme sera complètement démontré.

14.33. Revenons à l'égalité 14.31(2). Nous sommes maintenant en mesure de démontrer le théorème suivant:

Théorème. *Si une fonction $f(t)$ est continue par morceaux sur l'intervalle $[-\pi, \pi]$ et si la fonction $\frac{f(t_0+h)-f(t_0)}{h}$ est absolument intégrable par rapport à h au sens impropre dans un voisinage du point $h = 0$, alors les sommes partielles $s_{m,n}(t)$ de la série de Fourier de la fonction $f(t)$ convergent au point $t = t_0$ vers la valeur $f(t_0)$ lorsque $m \to \infty$ et $n \to \infty$ (indépendamment l'un de l'autre).*

Démonstration. Si la fonction $\frac{f(t_0+h)-f(t_0)}{h}$ est absolument intégrable au sens impropre dans un voisinage du point

$h = 0$, il en est de même de la fonction

$$\frac{f(t_0+h)-f(t_0)}{\sin\frac{h}{2}} = \frac{f(t_0+h)-f(t_0)}{h}\,\frac{h}{\sin\frac{h}{2}}.$$

Donc, en vertu du lemme, l'intégrale 14.31(2)

$$\frac{1}{4\pi i}\int_{-\pi}^{\pi}\frac{f(t_0+h)-f(t_0)}{\sin\frac{h}{2}}\left\{e^{i\left(n+\frac{1}{2}\right)h}-e^{-i\left(m+\frac{1}{2}\right)h}\right\}dh = \frac{1}{4\pi i}\left\{\int_{-\pi}^{0}+\int_{0}^{\pi}\right\}$$

tend vers zéro lorsque $h \to \infty$. Le théorème est démontré.

14.34. La condition d'intégrabilité absolue de $\frac{f(t_0+h)-f(t)}{h}$ par rapport à h pour $h \to 0$ s'appelle *condition de Dini* pour $f(t)$. Elle est satisfaite, par exemple, si la fonction $f(t)$ vérifie la *condition de Lipschitz d'ordre* $\alpha > 0$:

$$|f(t+h)-f(t)| \leqslant C\,|h|^{\alpha}.$$

En particulier, lorsque la fonction $f(t)$ possède au point t_0 une dérivée finie, la condition de Lipschitz d'ordre 1 est vérifiée et, par conséquent, les nombres $s_{m,n}(t_0)$ convergent vers $f(t_0)$.

14.35. Convergence uniforme d'une série de Fourier sur un ensemble $E \subset Q$. L'analyse de la démonstration ci-dessus montre que l'on peut obtenir, dans la même voie, la convergence uniforme de la série de Fourier sur un ensemble $E \subset Q$.

Nous dirons que la condition de Dini pour la fonction $f(t)$ est remplie *uniformément sur un ensemble* $E \subset Q$ si, pour tout $\varepsilon > 0$, on peut trouver un $\delta > 0$ tel que l'inégalité

$$\int_{|h|<\delta}\left|\frac{f(t+h)-f(t)}{h}\right|dh < \varepsilon$$

ait lieu pour tous les $t \in E$ à la fois.

Théorème. *Si une fonction $f(t)$ est bornée et continue par morceaux sur l'intervalle $[-\pi, \pi] = Q$ (les points $-\pi$ et π étant comme toujours identifiés) et si la condition de Dini pour $f(t)$ est satisfaite uniformément sur un ensemble $E \subset Q$, alors la série de Fourier de la fonction $f(t)$ converge vers celle-ci uniformément sur l'ensemble E.*

Démonstration. Nous avons déjà mis la différence $s_{m,n}(t) - f(t)$ sous la forme

$$s_{m,n}(t)-f(t) = \frac{1}{4\pi i}\int_{-\pi}^{\pi}|f(t+h)-f(t)|\,\frac{e^{i\left(n+\frac{1}{2}\right)h}-e^{-i\left(m+\frac{1}{2}\right)h}}{\sin\frac{h}{2}}\,dh.$$

Prouvons que cette différence tend vers zéro uniformément sur l'ensemble E. Posons $\varphi(t, h) = \dfrac{f(t+h)-f(t)}{\sin\frac{h}{2}}$. Par hypothèse, pour un $\varepsilon > 0$ donné, il existe un $\delta_0 > 0$, qui ne dépend pas de t, tel que

$$\int\limits_{|h|<\delta_0} \left| \frac{f(t+h)-f(t)}{\sin\frac{h}{2}} \right| dh = \int\limits_{|h|<\delta_0} \left| \frac{f(t+h)-f(t)}{h} \right| \frac{h}{\sin\frac{h}{2}}\, dh \leqslant \frac{\varepsilon}{3}.$$

A l'extérieur de l'intervalle $|h| \leqslant \delta_0$, la fonction $\varphi(t, h)$ est bornée en module par le nombre $M(\varepsilon) = \dfrac{2M}{\sin\frac{\delta_0}{2}}$, où $M = \max |f(t)|$; l'estimation donnée ne dépend pas de t. Mais l'ensemble Z_t des points de discontinuité de la fonction $\varphi(t, h)$ en dépend bien; notamment il s'obtient en translatant de $-t$ l'ensemble des points de discontinuité de la fonction $f(h)$ (sans compter le point de discontinuité éventuel $h = 0$ que nous avons déjà isolé). Par conséquent, l'ensemble Z_t, pour tout t, peut être recouvert par un translaté convenable S_t d'un système fini fixe d'intervalles dont la somme des longueurs est $\leqslant \dfrac{\varepsilon}{3M(\varepsilon)}$. C'est l'ensemble G_t, réunion d'un nombre fixe $N(\varepsilon)$ d'intervalles sur lesquels la fonction $\varphi(t, h)$ est continue (leur nombre ne peut que diminuer car certains d'eux se trouvent dans l'intervalle déjà isolé $|h| \leqslant \delta_0$), qui reste à l'extérieur de S_t.

Soit

$$g(h) = \frac{1}{\sin\frac{h}{2}} \qquad (|h| \geqslant \delta_0).$$

L'oscillation $\omega_\varphi(\delta)$ de la fonction $\varphi(t, h)$ peut être estimée, d'après 5.17*d*, comme suit:

$$\omega_\varphi(\delta) \leqslant \max |f(t+h) - f(t)|\, \omega_g(\delta) + \omega_f(\delta) \max |g(h)| \leqslant$$
$$\leqslant 2M_f\omega_g(\delta) + M_g\omega_f(\delta).$$

Maintenant l'estimation 14.32(5) fournit

$$4\pi |s_{m,n}(t) - f(t)| = \left| \int\limits_{-\pi}^{\pi} \varphi(h, t) \left[e^{i\left(n+\frac{1}{2}\right)h} - e^{-i\left(m+\frac{1}{2}\right)h}\right] dh \right| \leqslant$$

$$\leqslant \frac{\varepsilon}{3} + \left[2M_f\omega_g\left(\frac{2\pi}{m+\frac{1}{2}}\right) + M_g\omega_f\left(\frac{2\pi}{m+\frac{1}{2}}\right)\right] 2\pi +$$

$$+ N(\varepsilon) M(\varepsilon) \frac{2\pi}{m+\frac{1}{2}} + \frac{\varepsilon}{3} +$$

$$+ \left[2M_f\omega_g\left(\frac{2\pi}{n+\frac{1}{2}}\right) + M_g\omega_f\left(\frac{2\pi}{n+\frac{1}{2}}\right)\right] 2\pi + N(\varepsilon) M(\varepsilon) \frac{2\pi}{n+\frac{1}{2}}.$$

Si n et m sont suffisamment grands, le dernier membre devient inférieur à ε pour tous les $t \in E$ à la fois, ce qu'il nous faut. Le théorème est démontré.

14.36. Conséquence. *Si la condition de Lipschitz d'ordre* $\alpha > 0$

$$| f(t+h) - f(t) | \leqslant C | h |^{\alpha}$$

est satisfaite pour tout point t d'un ensemble $E \subset Q$ et que la constante C ne dépende pas du point $t \in E$, alors la série de Fourier de la fonction $f(t)$ converge vers elle uniformément sur E. En particulier, si la fonction $f(t)$ possède une dérivée bornée sur un intervalle $[c, d] \subset Q$ (respectivement dérivée à droite ou à gauche aux points c et d), alors, quel que soit un intervalle intérieur $[c+\delta, d-\delta]$, la condition de Lipschitz d'ordre 1 y est remplie pour tout $| h | \leqslant \delta$:

$$| f(t+h) - f(t) | \leqslant | h | \sup_{t \in [c,\, d]} | f'(t) |;$$

par conséquent, la série de Fourier de la fonction $f(t)$ converge vers celle-ci uniformément sur tout intervalle $[c+\delta, d-\delta]$.

14.37. Si la condition de Dini n'est pas satisfaite en un point, le théorème 14.33 n'est plus valable, et la série de Fourier de la fonction $f(t)$ peut s'avérer divergente (nous le verrons dans 14.51). La relation

$$\lim_{\substack{m\to\infty \\ n\to\infty}} s_{m,\,n}(t_0) = f(t_0)$$

ne peut être vérifiée qu'au sens d'un passage généralisé à la limite. Il est naturel de considérer, avant tout, les sommes partielles symétriques

$$s_{n,\,n}(t) = \sum_{k=-n}^{n} c_k e^{ikt}.$$

Pour une somme partielle symétrique $s_{n,\,n}(t)$ (qui se notera par la suite $s_n(t)$ tout court), on obtient de 14.31(1) l'expression

suivante:

$$s_n(t)=\frac{1}{4\pi i}\int_{-\pi}^{\pi} f(t+h)\,\frac{e^{i\left(n+\frac{1}{2}\right)h}-e^{-i\left(n+\frac{1}{2}\right)h}}{\sin\frac{h}{2}}\,dh=$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t+h)\,\frac{\sin\left(n+\frac{1}{2}\right)h}{\sin\frac{h}{2}}\,dh=\int_{-\pi}^{\pi} f(t+h)\,D_n(h)\,dh, \quad (1)$$

où

$$D_n(h)=\frac{1}{2\pi}\,\frac{\sin\left(n+\frac{1}{2}\right)h}{\sin\frac{h}{2}}$$

est le *noyau de Dirichlet*. Si les fonctions $D_n(h)$ formaient une suite en forme de delta pour le point 0, on pourrait appliquer le théorème 12.55*b* et obtenir directement la convergence des sommes symétriques de la série de Fourier de la fonction $f(t)$ vers sa valeur $f(t_0)$ en tout point t_0 de continuité de $f(t)$. Mais les fonctions $D_n(h)$ ne constituent pas une suite en forme de delta, ce que nous verrons dans 14.51. La fonction $D_n(h)$ est évidemment paire: $D_n(-h)=D_n(h)$; de plus, en posant $f(t)\equiv 1$ dans (1), nous trouvons $s_n(t)\equiv 1$ et

$$\int_{-\pi}^{\pi} D_n(h)\,dh=1.$$

En nous servant de ces propriétés nous allons étudier la convergence des sommes symétriques de la série de Fourier de la fonction $f(t)$ en ses points de discontinuité de première espèce.

14.38. Comportement de la série de Fourier aux points de discontinuité de 1-ère espèce de la fonction $f(t)$. Soit t_0 un point de discontinuité de 1-ère espèce de la fonction $f(t)$, de sorte qu'il existe les valeurs

$$f(t_0+0)=\lim_{t\searrow t_0} f(t), \qquad f(t_0-0)=\lim_{t\nearrow t_0} f(t).$$

Supposons que *les conditions unilatérales de Dini* soient satisfaites, i.e. que les intégrales

$$\int_{t_0}^{t_0+\delta}\left|\frac{f(t)-f(t_0+0)}{t}\right|dt, \qquad \int_{t_0-\delta}^{t_0}\left|\frac{f(t)-f(t_0-0)}{t}\right|dt$$

convergent pour un $\delta>0$. Tout comme avant, la quantité $s_n(t_0)$ peut s'écrire:

$$s_n(t_0)=\int_{-\pi}^{\pi}[f(t_0+h)]\,D_n(h)\,dh.$$

En introduisant les valeurs limites $f(t_0 + 0)$ et $f(t_0 - 0)$, transformons l'expression obtenue :

$$s_n(t_0) = \int_{-\pi}^{0} |f(t_0 + h) - f(t_0 - 0)| D_n(h)\, dh +$$

$$+ \int_{0}^{\pi} |f(t_0 + h) - f(t_0 + 0)| D_n(h)\, dh +$$

$$+ f(t_0 - 0) \int_{-\pi}^{0} D_n(h)\, dh + f(t_0 + 0) \int_{0}^{\pi} D_n(h)\, dh =$$

$$= I_1 + I_2 + |f(t_0 - 0) + f(t_0 + 0)| \int_{0}^{\pi} D_n(h)\, dh,$$

où l'on a utilisé la parité du noyau de Dirichlet $D_n(h)$. En vertu du lemme 14.32, les quantités I_1 et I_2 tendent vers zéro lorsque $n \to \infty$. Le dernier terme ne dépend pas de n et vaut $\frac{f(t_0 + 0) + f(t_0 - 0)}{2}$. Ainsi, *lorsque les conditions unilatérales de Dini sont satisfaites en un point de discontinuité de* 1-ère *espèce*, par exemple s'il y a les dérivées

$$f'(t_0 + 0) = \lim_{h \searrow 0} \frac{f(t_0 + h) - f(t_0 + 0)}{h},$$

$$f'(t_0 - 0) = \lim_{h \searrow 0} \frac{f(t_0 - 0) - f(t_0 - h)}{h},$$

les sommes partielles symétriques de la série de Fourier de la fonction $f(t)$ *convergent vers le nombre* $\frac{1}{2}|f(t_0 + 0) + f(t_0 - 0)|$.

Si l'on met la série de Fourier de la fonction $f(t)$ sous la forme

$$\frac{a_0}{2} + \sum_{1}^{\infty} (a_k \cos kt + b_k \sin kt), \tag{1}$$

alors, comme nous l'avons déjà vu dans 14.23, la n-ième somme partielle de cette série

$$\frac{a_0}{2} + \sum_{1}^{n} (a_k \cos kt + b_k \sin kt) \tag{2}$$

coïncide avec la n-ième somme symétrique de la série de Fourier de la fonction $f(t)$ sous forme complexe ; donc, pour la série (1), l'énoncé du théorème sur la convergence est le même pour les points de continuité (où la condition bilatérale de Dini est satisfaite) et

pour les points de discontinuité de 1-ère espèce (où les conditions unilatérales de Dini sont satisfaites).

Notons encore que, puisque $a_n \to 0$, $b_n \to 0$ (14.26*a*), le groupement des termes dans la série (1) indiqué par les parenthèses n a aucune importance pour la convergence ou la divergence de la série.

§ 14.4. Autres propriétés des séries de Fourier. Applications

14.41. Calcul des coefficients de Fourier. Signalons quelques propriétés simples des coefficients de Fourier qui facilitent leur calcul.

a. Si $f(t)$ est une fonction paire, i.e. $f(-t) = f(t)$, alors on a

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt = 0,$$

et $f(t)$ se développe en une série de Fourier suivant les fonctions $\cos nt$. De plus,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt = \frac{2}{\pi} \int_{0}^{\pi} f(t) \cos nt \, dt. \tag{1}$$

b. Si $f(t)$ est une fonction impaire, i.e. $f(-t) = -f(t)$, alors on a

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt = 0,$$

et $f(t)$ se développe en une série de Fourier suivant les fonctions $\sin nt$. De plus,

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt = \frac{2}{\pi} \int_{0}^{\pi} f(t) \sin nt \, dt. \tag{2}$$

14.42. a. Soit $f(t)$ une fonction polynomiale par morceaux, c'est-à-dire que l'intervalle $[-\pi, \pi]$ se décompose en un nombre fini d'intervalles $[t_j, t_{j+1}]$, $j = 0, 1, \ldots, m-1$ sans points intérieurs communs, de sorte que sur l'intervalle $[t_j, t_{j+1}]$ la fonction coïncide avec un polynôme $p_j(t) = \sum_{k=0}^{s} p_{jk} t^k$. Alors on a

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} \, dt = \frac{1}{2\pi} \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} p_j(t) e^{-int} \, dt.$$

Transformons chaque terme en intégrant par parties :

$$\int_{t_j}^{t_{j+1}} p_j(t)\, e^{-int}\, dt = p_j(t)\, \frac{e^{-int}}{-in}\Big|_{t_j}^{t_{j+1}} - \int_{t_j}^{t_{j+1}} p_j'(t)\, \frac{e^{-int}}{-in}\, dt =$$

$$= \frac{1}{in}\left[p_j(t_j)\, e^{-int_j} - p_j(t_{j+1})\, e^{-int_{j+1}}\right] - \frac{i}{n}\int_{t_j}^{t_{j+1}} p_j'(t)\, e^{-int}\, dt.$$

En intégrant un nombre fini de fois, on obtiendra une expression sans intégrales, et les coefficients de Fourier auront la forme de polynômes en $1/n$ et e^{int_j}. Le calcul analogue pour les coefficients a_n et b_n montre qu'ils sont des polynômes en $1/n$, $\cos nt_j$ et $\sin nt_j$.

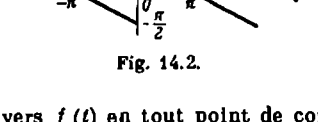


Fig. 14.2.

Il résulte des théorèmes généraux du § 14.3 que la série de Fourier de toute fonction polynomiale par morceaux $f(t)$ converge vers $f(t)$ en tout point de continuité de $f(t)$. Cette convergence est uniforme sur chaque intervalle qui ne contient pas à son intérieur ni sur sa frontière de points de discontinuité de la fonction $f(t)$. En tout point de discontinuité, les sommes symétriques de la série de Fourier convergent vers la valeur $\frac{1}{2}|f(t+0) + f(t-0)|$.

b. Exemple. Considérons la fonction $f(t) = \frac{\pi - t}{2}$ $(0 < t < \pi)$ prolongée d'une façon impaire sur l'intervalle $-\pi < t < 0$, puis, d'une façon périodique de période 2π, sur tout l'axe $-\infty < t < \infty$ (fig. 14.2). Conformément à 14.41*b*, la fonction $f(t)$ se développe en sa série de Fourier suivant les fonctions $\sin nt$, de sorte que

$$\frac{\pi}{2}\, b_n = \int_0^{\pi} \frac{\pi - t}{2} \sin nt\, dt = \frac{\pi - t}{2}\, \frac{\cos nt}{n}\Big|_{\pi}^{0} + \frac{1}{2}\int_{\pi}^{0} \frac{\cos nt}{n}\, dt = \frac{\pi}{2n}.$$

Ainsi, sur $(0, 2\pi)$, nous avons

$$\frac{\pi - t}{2} = \sum_{n=1}^{\infty} \frac{\sin nt}{n},$$

la série étant convergente en tout point $t \in (0, 2\pi)$ uniformément dans tout intervalle $[\delta, 2\pi - \delta]$, $\delta > 0$. Aux points $t = 0$ et $t = 2\pi$, la somme de la série est nulle en vertu de 14.38.

c. Dans certains cas, étant donnés les coefficients a_n et b_n sous forme de polynômes en $1/n$, $\cos nt_j$ et $\sin nt_j$, on peut sommer la série de Fourier et obtenir une formule explicite pour la fonction

polynomiale par morceaux $f(t)$ [12]. Cependant, même une série si simple que $\sum_{1}^{\infty} \frac{\cos nt}{n}$ n'est pas la série de Fourier d'une fonction polynomiale par morceaux (cf. 14.47).

14.43. Liaison entre la dérivabilité de la fonction $f(t)$ et l'ordre de décroissance de ses coefficients de Fourier.

a. Soit $f(t)$ une fonction continue sur la circonférence $Q = [-\pi, \pi]$ et possédant la dérivée continue par morceaux $f'(t)$. Soient c_n les coefficients de Fourier de la fonction $f(t)$ (par rapport au système e^{int}) et c'_n les coefficients de Fourier de la fonction $f'(t)$. Nous avons

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt =$$

$$= \frac{1}{2\pi} f(t) \frac{e^{-int}}{-in} \Big|_{-\pi}^{\pi} + \frac{1}{2\pi in} \int_{-\pi}^{\pi} f'(t) e^{-int} dt = \frac{c'_n}{in}. \qquad (1)$$

Le terme sans intégrale est ici nul, car $f(-\pi) = f(\pi)$ d'après l'hypothèse de continuité de la fonction $f(t)$ sur toute la circonférence Q. Les nombres c'_n en tant que coefficients de Fourier d'une fonction continue par morceaux tendent vers zéro; nous voyons que les coefficients de Fourier d'une fonction $f(t)$ dérivable tendent vers zéro plus rapidement que $1/n$. En outre, la série des nombres $|c_n|$ converge, ce qui découle de l'inégalité

$$|c_n| = \frac{1}{|n|} |c'_n| \leqslant \frac{1}{2} \left(\frac{1}{n^2} + |c'_n|^2 \right)$$

et de la convergence de la série des nombres $|c'_n|^2$. (Vu le critère de Weierstrass 6.53, cela démontre, sans que le théorème 14.36 soit utilisé, la convergence uniforme de la série de Fourier pour une fonction $f(t)$ qui satisfait aux conditions imposées dans ce numéro.) Si la fonction $f(t)$ est continue mais l'existence de sa dérivée n'est pas supposée, alors la convergence de la série des nombres $|c_n|$ n'a en général pas lieu, ce que nous verrons plus bas (14.53).

b. Si la fonction $f(t)$ est continue et possède les dérivées continues y compris la $(m-1)$-ième et que $f^{(m)}(t)$ soit continue par morceaux, alors on peut poursuivre la transformation (1) en désignant par $c_n^{(k)}$ les coefficients de Fourier de la fonction $f^{(k)}(t)$:

$$c_n = \frac{c'_n}{in} = \frac{c''_n}{(in)^2} = \ldots = \frac{c_n^{(m-1)}}{(in)^{m-1}} = \frac{c_n^{(m)}}{(in)^m} \quad (n = \pm 1, \pm 2, \ldots). \qquad (2)$$

Dans ce cas, les coefficients $c_n^{(m-1)}$ forment une série absolument convergente (cf. a); donc, en plus des égalités (2), on peut écrire

l'expression suivante pour le coefficient c_n :

$$c_n = \frac{\varepsilon_n}{|n|^{m-1}} \quad (n = \pm 1, \pm 2 \ldots),$$

où la série $\sum_{-\infty}^{\infty} |\varepsilon_n|$ converge.

14.44.a. Les propositions précédentes peuvent être inversées, ne serait-ce que partiellement. Supposons que les coefficients de Fourier c_n d'une fonction $f(t)$ aient la forme

$$c_n = \frac{\theta_n}{|n|^m}, \quad |\theta_n| \leqslant c, \quad m \geqslant 2,$$

ou

$$c_n = \frac{\varepsilon_n}{|n|^{m-2}}, \quad \sum_{-\infty}^{\infty} |\varepsilon_n| < \infty, \quad m \geqslant 2.$$

Alors, $f(t)$ *a les dérivées continues y compris la* $(m-2)$-ième. En effet, dans l'hypothèse formulée, la série de Fourier de la fonction $f(t)$ est uniformément convergente (d'après le critère de Weierstrass) de même que les séries que l'on obtient par les dérivations successives formelles jusqu'à l'ordre $m-2$:

$$\sum_{-\infty}^{\infty} c_n e^{int} \equiv s_0(t),$$

$$\sum_{-\infty}^{\infty} c_n (in) e^{int} \equiv s_1(t),$$

. . . . . .

$$\sum_{-\infty}^{\infty} c_n (in)^{m-2} e^{int} \equiv s_{m-2}(t).$$

En vertu de 14.26*c*, la fonction $f(t)$ est identique à $s_0(t)$. D'après la théorème 9.78 sur la dérivation d'une suite de fonctions, la fonction $s_0(t)$ est dérivable, et sa dérivée coïncide avec $s_1(t)$; le fait analogue a lieu pour la fonction $s_1(t)$, etc. ; définitivement, la fonction $f(t)$ s'avère $m-2$ fois continûment dérivable.

b. Si la fonction $f(t)$ a les dérivées continues de tous les ordres $m = 1, 2, \ldots$, alors ses coefficients de Fourier vérifient les inégalités

$$|c_n| \leqslant \frac{\theta_m}{|n|^m}, \quad m = 1, 2, \ldots, \tag{1}$$

donc décroissent plus rapidement que n'importe quelle puissance de $1/|n|$. Inversement, si les coefficients de Fourier d'une fonction $f(t)$ satisfont aux inégalités (1) pour tout $m = 1, 2, \ldots,$

alors, d'après ce qu'on a dit plus haut, la fonction $f(t)$ est continue et a les dérivées continues de tous les ordres. Ainsi, la classe des fonctions indéfiniment dérivables est complètement caractérisée par les conditions (1) imposées aux coefficients de Fourier c_n.

14.45*. Problème des isopérimètres. On appelle ainsi le problème classique suivant : de toutes les courbes planes fermées lisses par morceaux de longueur donnée, trouver celle qui entoure l'aire maximale. C'est la circonférence qui en est la solution ; pour le prouver nous effectuons, avec Hurwitz, la construction suivante. Soit $z(s) = x(s) + iy(s)$ la représentation paramétrique d'une courbe plane fermée lisse par morceaux L, avec la longueur d'arc s pour paramètre (9.63g). Supposons d'abord que la longueur totale de la courbe L soit 2π, de sorte que $z(2\pi) = z(0)$. Ecrivons les développements de Fourier des fonctions $x(s)$ et $y(s)$:

$$x(s) = \frac{a_0}{2} + \sum_1^\infty (a_n \cos ns + b_n \sin ns), \tag{1}$$

$$y(s) = \frac{c_0}{2} + \sum_1^\infty (c_n \cos ns + d_n \sin ns) ; \tag{2}$$

d'après 14.43a nous avons

$$x'(s) = \sum_1^\infty (-na_n \sin ns + nb_n \cos ns),$$

$$y'(s) = \sum_1^\infty (-nc_n \sin ns + nd_n \cos ns).$$

Comme $|x'(s)|^2 + |y'(s)|^2 = 1$ (9.63g), nous trouvons, à l'aide de 14.25(1) :

$$2\pi = \int_0^{2\pi} \{[x'(s)]^2 + [y'(s)]^2\}\, ds =$$

$$= \pi \sum_1^\infty n^2 (a_n^2 + b_n^2 + c_n^2 + d_n^2). \tag{3}$$

D'autre part, en appliquant la formule 9.94(4) pour l'aire G intérieure à une courbe fermée et compte tenu de 14.25(3), nous trouvons

$$G = \frac{1}{2} \int_0^{2\pi} [xy'(s) - yx'(s)]\, ds = \pi \sum_{n=1}^\infty n (a_n d_n - b_n c_n). \tag{4}$$

A partir des égalités (3) et (4), on a

$$2-\frac{2G}{\pi}=\sum_1^\infty \{n^2(a_n^2+b_n^2+c_n^2+d_n^2)-2n(a_nd_n-b_nc_n)\}=$$

$$=\sum_1^\infty \{(na_n-d_n)^2+(nb_n+c_n)^2+(n^2-1)(c_n^2+d_n^2)\}\geqslant 0. \quad (5)$$

Ainsi, *l'aire G intérieure à une courbe fermée de longueur* 2π *ne dépasse pas* π. Examinons le cas d'égalité dans (5) ; pour tout $n = 1, 2, \ldots$, on a alors

$$na_n-d_n=0,\ nb_n+c_n=0,\ (n^2-1)(c_n^2+d_n^2)=0.$$

En particulier, si $n > 1$, on obtient $c_n = d_n = 0$, d'où $a_n = b_n = 0$ pour les mêmes $n > 1$. En posant $n = 1$ on trouve $a_n = d_n$, $b_n = -c_n$. Il découle alors de (3) que $a_1^2 + b_1^2 = 1$, et l'on peut poser $a_1 = \cos\alpha$, $b_1 = \sin\alpha$. En portant dans (1) et (2) on obtient définitivement

$$x(s)=\frac{a_0}{2}+\cos(s-\alpha),$$

$$y(s)=\frac{c_0}{2}+\sin(s-\alpha);$$

cette courbe est la circonférence de rayon 1 centrée au point $(a_0/2, c_0/2)$.

Si la longueur totale de la courbe L est égale à un nombre $l \neq 2\pi$, nous effectuons l'homothétie $x' = 2\pi x/l$, $y' = 2\pi y/l$ qui transforme la courbe L en une courbe L' de longueur 2π (9.63*e*) limitant l'aire $G' = (2\pi/l)^2 G$ (9.61*d*, *e*). D'après ce qu'on a vu,

$$G'\leqslant\pi,\quad G\leqslant\frac{l^2}{4\pi},$$

et dans le cas extrême où L' est une circonférence de rayon 1, la courbe L est également une circonférence mais de rayon $l/(2\pi)$.

14.46. Emploi d'une variable complexe. Désignons les points de la circonférence unité $Q = \{x^2 + y^2 = 1\}$ à l'aide de la variable complexe $z = x + iy = e^{it}$, $-\pi \leqslant t \leqslant \pi$. Posons $f(t) \equiv F(z)$. La série de Fourier de la fonction $f(t)$ prend la forme

$$f(t)\equiv F(z)=\sum_{-\infty}^{\infty}c_ne^{int}=\sum_{-\infty}^{\infty}c_nz^n. \quad (1)$$

Nous avons déjà rencontré dans 10.45 une série de la forme (1) suivant les puissances de z (série de Laurent). Les coefficients c_n exprimés par les intégrales par rapport à t peuvent également être exprimés

par les intégrales par rapport à la variable complexe z si l'on utilise l'égalité

$$dz = ie^{it}\,dt = iz\,dt.$$

En effet,

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,e^{-int}\,dt = \frac{1}{2\pi i}\oint_{|z|=1} F(z)\,z^{-n-1}\,dz. \qquad (2)$$

Si la fonction $F(z)$ est analytiquement prolongeable à l'intérieur du cercle unité, alors la série de Laurent, donc aussi celle de Fourier (1), devient série de Taylor

$$F(z) = \sum_0^\infty c_n z^n.$$

14.47. Exemple. Calculons la somme de la série $\sum_1^\infty \frac{\cos nt}{n}$.

On peut écrire

$$\sum_1^\infty \frac{\cos nt}{n} + i\sum_1^\infty \frac{\sin nt}{n} = \sum_1^\infty \frac{e^{int}}{n},$$

ce qui ramène le problème au calcul de la somme $\sum_1^\infty \frac{e^{int}}{n}$ ou de la somme $\sum_1^\infty \frac{z^n}{n}$. La dernière série se déduit en intégrant terme à terme, de 0 à z, la série

$$\sum_0^\infty \zeta^n = \frac{1}{1-\zeta} \qquad (1)$$

On sait (10.57) que

$$\int_0^z \frac{d\zeta}{1-\zeta} = -\int_1^{1-z} \frac{d\omega}{\omega} = -\ln(1-z),$$

si l'on choisit pour chemin d'intégration dans le plan des ω une courbe qui ne coupe pas le demi-axe réel négatif, ce qui correspond, dans le plan des ζ ($1 - \zeta = \omega$), à une courbe qui ne coupe pas la partie $\zeta \geqslant 1$ du demi-axe réel positif. Il nous suffit d'intégrer le long des segments de droite $[0, z[$ (ce qui assure, en particulier, l'univocité de la fonction $-\ln(1 - z)$). La série (1) converge pour

$|\zeta| < 1$, de sorte que

$$-\ln(1-z) = \sum_1^\infty \frac{z^n}{n} \qquad (2)$$

pour $|z| < 1$. Or, la série (2) converge également pour $|z| = 1$, $z \neq 1$ (6.63c); comme la fonction $\ln(1-z)$ reste continue en ces points, l'égalité (2) reste valable pour eux, d'après le théorème d'Abel 6.67. Pour tout $w = |w| e^{i \arg w}$, on a (10.57)

$$\ln w = \ln |w| + i \arg w.$$

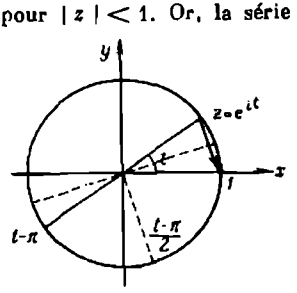


Fig. 14.3.

En particulier, si $z = e^{it}$, $\operatorname{Im} z > 0$, le module et l'argument de la quantité $1 - z$ sont facilement déterminés d'après la fig. 14.3. Notamment,

$$|1-z| = 2\sin\frac{t}{2}, \quad \arg(1-z) = \frac{t-\pi}{2},$$

ensuite

$$\ln|1-z| = \ln\left(2\sin\frac{t}{2}\right),$$

$$\ln(1-z) = \ln\left(2\sin\frac{t}{2}\right) + i\,\frac{t-\pi}{2},$$

d'où, pour $t \in (0, 2\pi)$,

$$\sum_1^\infty \frac{z^n}{n} = -\ln(1-z) = -\ln\left(2\sin\frac{t}{2}\right) + i\,\frac{\pi-t}{2},$$

$$\sum_1^\infty \frac{\cos nt}{n} = -\ln\left(2\sin\frac{t}{2}\right), \quad \sum_1^\infty \frac{\sin nt}{n} = \frac{\pi-t}{2}.$$

La dernière égalité nous est déjà familière (14.42*b*).

14.48*. Problème de solutions périodiques. Soit

$$a_0 u^{(m)}(t) + a_1 u^{(m-1)}(t) + \ldots + a_m u(t) = g(t) \qquad (1)$$

une équation différentielle linéaire à coefficients constants dont le second membre $g(t)$ est périodique de période 2π. On se demande si cette équation a une solution $u(t)$ de période 2π.

Cherchons une telle solution sous forme de série de Fourier

$$u(t) = \sum_{-\infty}^{\infty} u_k e^{ikt}, \qquad (2)$$

avec les coefficients inconnus u_k. En supposant que m dérivations successives terme à terme de la série (2) soient légitimes et en désignant $a_0\lambda^m + a_1\lambda^{m-1} + \ldots + a_m \equiv p(\lambda)$ nous avons

$$\sum_{-\infty}^{\infty} u_k p(ik) e^{ikt} = a_0 u^{(m)}(t) + \ldots + a_m u(t). \tag{3}$$

D'autre part, soit

$$g(t) = \sum_{-\infty}^{\infty} g_k e^{ikt} \tag{4}$$

le développement de Fourier de la fonction $g(t)$. En comparant les développements orthogonaux (3) et (4) nous trouvons, pour tout $k = 0, \pm 1, \pm 2, \ldots$

$$g_k = p(ik)\, u_k,$$

d'où, pour $p(ik) \neq 0$,

$$u_k = \frac{g_k}{p(ik)},$$

donc la solution cherchée est

$$u(t) = \sum_{-\infty}^{\infty} \frac{g_k}{p(ik)} e^{ikt}. \tag{5}$$

Ces raisonnements euristiques conduisent aux énoncés suivants:

a. *Supposons que les coefficients de Fourier de la fonction périodique $g(t)$ forment une série absolument convergente (par exemple, $g(t)$ est lisse par morceaux). Si $p(ik)$ ne s'annule pas pour $k = 0, \pm 1, \pm 2, \ldots$, alors l'équation (1) possède une solution périodique et une seule.*

En effet, dans l'hypothèse formulée, le polynôme $p(ik)$ de degré m admet la minoration

$$|p(0)| \geqslant c, \quad |p(ik)| \geqslant c\,|k|^m \quad (k = \pm 1, \pm 2, \ldots).$$

Il en résulte que

$$\left|\frac{g_k}{p(ik)}\right| \leqslant \frac{|g_k|}{c\,|k|^m}.$$

En vertu de 14.44*a*, la fonction $u(t)$ définie par l'égalité (5) possède les dérivées continues y compris la m-ième, et ces dérivées peuvent être obtenues en dérivant terme à terme la série (5). En les portant dans l'équation (1), cette dernière se trouve satisfaite. Si, en plus de la solution périodique trouvée, il en existe encore une, leur différence $v(t)$ est une solution périodique de l'équation

$$a_0 v^{(m)}(t) + a_1 v^{(m-1)}(t) + \ldots + a_m v(t) = 0. \tag{6}$$

Or, nous connaissons la solution générale de l'équation (6) qui s'exprime par les fonctions de la forme $e^{\lambda x}$, où $p(\lambda) = 0$ (13.18). Ces exponentielles ne conduisent à des solutions 2π-périodiques que si $\lambda = ik$ ($k = 0, \pm 1, \pm 2, \ldots$). Comme, par hypothèse, aucune des égalités écrites n'a lieu, il n'existe pas de solution périodique non nulle de l'équation (6). Cela démontre l'unicité de la solution périodique de l'équation (1).

b. *Si $p(ik) = 0$ pour certains $k = k_1, \ldots, k_r$ entiers, alors l'équation* (1) *possède une solution 2π-périodique si, et seulement si, $g_{k_j} = 0$ ($j = 1, \ldots, r$) (l'hypothèse du théorème précédent sur la convergence absolue de la série des nombres g_k restant valable); cette solution est définie au terme additif $\sum_{j=1}^{r} c_j e^{ik_j x}$ près, c_j étant des constantes arbitraires.*

En effet, si $g_{k_j} = 0$ ($j = 1, \ldots, r$), alors l'expression (5) avec les constantes arbitraires c_j pour coefficients $\frac{g_{k_j}}{p(ik_j)}$ (qui, dans le présent cas, ont la forme 0/0) représente de même que dans a une solution périodique de l'équation (1). Si, pour un certain $k = q$, nous avons $p(ik) = 0$, $g_q \neq 0$, alors, en portant ladite expression dans l'équation (1) et en multipliant scalairement l'identité obtenue par e^{-iqx}, nous voyons que $p(iq) = g_q = 0$. Cela signifie que l'équation (1) n'a aucune solution périodique. La démonstration de la dernière conclusion du théorème est analogue à celle de a.

14.49*. Parmi les applications nombreuses de séries de Fourier aux problèmes de physique mathématique, nous choisissons deux,

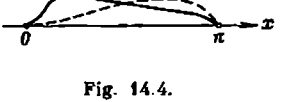


Fig. 14.4.

à savoir : résolution du problème de vibrations d'une corde homogène et celle du problème de figure d'équilibre d'une membrane circulaire.

a. Supposons qu'une corde soit fixée aux points 0 et π de l'axe des x et que, dans l'état d'équilibre, elle coïncide avec l'intervalle $[0, \pi]$ (fig. 14.4). Si l'on donne à la corde une forme quelconque représentée, par exemple, par une fonction $f(x)$ et qu'on la lâche ensuite, la corde se met à vibrer. Il faut trouver la fonction $u(t, x)$ déterminant la forme de la corde à l'instant t. L'équation pour la fonction $u(t, x)$ est déduite en physique mathématique ; dans certai-

nes conditions simplificatrices elle s'écrit

$$\frac{\partial^2 u(t, x)}{\partial t^2} = a^2 \frac{\partial^2 u(t, x)}{\partial x^2}, \tag{1}$$

où a est une constante. L'équation (1) est à résoudre avec les conditions initiales suivantes:

1) $u(0, x) = f(x)$ (la forme initiale est donnée);

2) $\frac{\partial u(0, x)}{\partial t} = 0$ (la vitesse initiale est nulle).

Nous allons résoudre le problème à l'aide de séries de Fourier. Notamment, développons la fonction $u(t, x)$ définie pour tout $t \geqslant 0$ fixe sur l'intervalle $[0, \pi]$ en sa série de Fourier suivant les fonctions $\sin nx$:

$$u(t, x) \backsim \sum_1^\infty b_n(t) \sin nx. \tag{2}$$

Les coefficients $b_n(t)$ sont à déterminer. Les conditions initiales 1)-2) sont bien satisfaites si l'on choisit les coefficients de façon que:

3) $b_n(0) = b_n$ est le coefficient de Fourier de la fonction $f(x)$;

4) $b_n'(0) = 0$.

Maintenant il faut que la fonction (2) satisfasse à l'équation (1). Formellement, nous avons

$$\frac{\partial^2 u(t, x)}{\partial t^2} = \sum_1^\infty b_n''(t) \sin nx, \tag{3}$$

$$\frac{\partial^2 u(t, x)}{\partial x^2} = -\sum_1^\infty n^2 b_n(t) \sin nx. \tag{4}$$

L'équation (1) est vérifiée si, pour tout $n = 1, 2, \ldots$, on a

$$b_n''(t) = -a^2 n^2 b_n(t). \tag{5}$$

La solution de l'équation (5) avec les conditions initiales 3)-4) est

$$b_n(t) = b_n \cos a\, nt, \tag{6}$$

et la solution (2) prend la forme définitive

$$u(t, x) = \sum_1^\infty b_n \cos a\, nt \sin nx. \tag{7}$$

Les dérivations formelles (3)-(4) sont légitimes si les séries correspondantes dans les deuxièmes membres convergent uniformément. Vu (6), il suffit pour cela que la série

$$\sum_1^\infty b_n n^2 \cos a\, nt \sin nx$$

converge uniformément sur l'intervalle $0 \leqslant x \leqslant \pi$; cela aura lieu, à son tour, si la série $\sum_1^\infty |b_n| n^2$ converge elle aussi. Or, la dernière série converge lorsque la fonction $f(x)$ est continue avec ses dérivées première et deuxième et si sa dérivée troisième est continue par morceaux (14.44*a*).

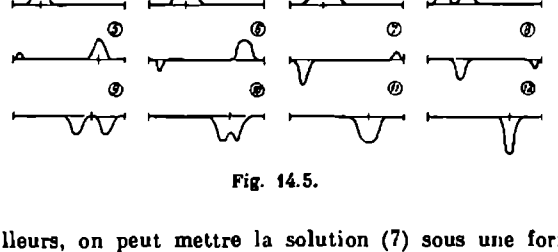


Fig. 14.5.

D'ailleurs, on peut mettre la solution (7) sous une forme qui n'exige aucune dérivabilité de la fonction $f(x)$. A savoir, nous avons

$$u(t, x) = \frac{1}{2} \sum_1^\infty b_n [\sin n(x+at) + \sin n(x-at)] =$$

$$= \frac{1}{2} [f(x+at) + f(x-at)]. \qquad (8)$$

Ici on sous-entend par $f(x + at)$ et $f(x - at)$, dans le cas où la variable $x + at$ (ou $x - at$) sort du domaine de définition initial $[0, \pi]$ de la fonction $f(x)$, le résultat du prolongement impair de $f(x)$ sur l'intervalle $[-\pi, 0]$ suivi du prolongement 2π-périodique sur tout l'axe $-\infty < x < \infty$.

Une question se pose : dans quel sens la fonction (8) vérifie-t-elle l'équation (1) lorsque la fonction $f(x)$ ne possède pas la dérivée deuxième? La physique mathématique vient sans grandes difficultés à bout de questions pareilles en généralisant la notion même de solution, mais nous n'insistons pas là-dessus [11]. La figure 14.5 montre les positions successives de la corde vibrante déterminées d'après la formule (8), la position initiale étant la première.

b. Figure d'équilibre d'une membrane circulaire. Supposons qu'une membrane soit disposée au-dessus du cercle $Q = \{x^2 + y^2 \leqslant 1\}$ et fixée par la circonférence $\Gamma = \{x^2 + y^2 = 1\}$ au-dessus de laquelle sa forme est donnée par une fonction continue $z = f(t)$ de l'angle polaire t. Dans l'état d'équilibre

(sous l'action des forces d'élasticité), la membrane prend la forme représentée par une fonction $z = u(x, y)$ (fig. 14.6). L'équation pour cette fonction est déduite en physique mathématique ; dans certaines conditions simplificatrices, elle est l'équation de Laplace

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \tag{9}$$

Il faut trouver une solution $u(x, y)$ de l'équation (9) continue dans tout le cercle Q (de telles fonctions sont dites harmoniques; cf. 10.18) et identique à la fonction donnée $f(t)$ sur la courbe Γ.

Pour le faire, développons la fonction $f(t)$ en sa série de Fourier :

$$f(t) \sim \frac{a_0}{2} + \sum_1^\infty (a_n \cos nt + b_n \sin nt).$$

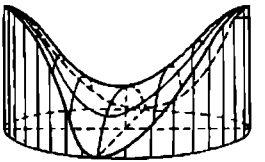


Fig. 14.6.

Nous mettons le symbole $\sim$ puisque la série de Fourier de la fonction donnée $f(t)$ peut ne pas converger vers celle-ci (nous le verrons dans 14.51). Posons en coordonnées polaires t, r:

$$u(r, t) = \frac{a_0}{2} + \sum_1^\infty r^n (a_n \cos nt + b_n \sin nt) \quad (r < 1). \tag{10}$$

La fonction $u(r, t)$ est la partie réelle de la fonction analytique

$$\frac{a_0}{2} + \sum_1^\infty (a_n - ib_n) z^n \quad (r = |z| < 1), \tag{11}$$

donc vérifie l'équation de Laplace (10.18) à l'intérieur du cercle Q. Montrons que les autres conditions sont également satisfaites. En explicitant les coefficients de Fourier nous avons

$$u(r, t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\tau)\, d\tau +$$

$$+ \frac{1}{\pi} \sum_1^\infty \int_{-\pi}^{\pi} f(\tau) (\cos nt \cos n\tau + \sin nt \sin n\tau) r^n\, d\tau =$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau) \left\{ \frac{1}{2} + \sum_1^\infty r^n \cos n(t - \tau) \right\} d\tau.$$

Ici on peut permuter la sommation et l'intégration puisque, d'après le critère de Weierstrass 6.53, la série entre accolades converge uniformément par rapport à t pour $r < 1$.

D'après 6.47*f*, nous avons

$$\frac{1}{2}+\sum_{1}^{\infty} r^n \cos n\theta = \frac{1-r\cos\theta}{1-2r\cos\theta+r^2}-\frac{1}{2}=\frac{1-r^2}{2(1-2r\cos\theta+r^2)},$$

d'où

$$u(r,\ t)=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\tau)\,\frac{1-r^2}{1-2r\cos(t-\tau)+r^2}\,d\tau=$$

$$=\int_{-\pi}^{\pi} f(\tau)\,P_r(t-\tau)\,d\tau, \qquad (12)$$

où

$$P_r(t)=\frac{1}{2\pi}\,\frac{1-r^2}{1-2r\cos t+r^2} \qquad (r<1);$$

cette fonction est appelée *noyau de Poisson.*

Comme le dénominateur du noyau de Poisson a la forme

$$1-2r\cos t+r^2=(1-r)^2+4r\sin^2\frac{t}{2},$$

le noyau de Poisson n'est pas négatif. Vérifions qu'il possède les propriétés d'une fonction de t en forme de delta pour $r\to 1$. En posant dans (12) $f(t)\equiv 1$ nous tirons de (10) $u(r,\ t)\equiv 1$, donc

$$\int_{-\pi}^{\pi} P_r(\tau)\,d\tau=1.$$

Ensuite, pour tout $\delta>0$, on a l'estimation

$$\int_{|t|\geqslant\delta} P_r(\tau)\,d\tau\leqslant\frac{1-r^2}{2\pi}\int_{|t|\geqslant\delta}\frac{d\tau}{(1-r)^2+4r\sin^2\frac{\tau}{2}}\leqslant\frac{1-r^2}{4r\sin^2\frac{\delta}{2}}$$

qui montre que

$$\lim_{r\to 1}\int_{|t|\geqslant\delta} P_r(\tau)\,d\tau=0.$$

En appliquant le théorème 12.55*g* sur les fonctions en forme de delta nous voyons que la fonction $u(r,\ t)$ complétée par les valeurs $f(t)$ sur la courbe Γ est continue dans le cercle Q, ce qu'il nous fallait.

Dans la théorie des équations aux dérivées partielles, on démontre l'unicité de la solution trouvée dans la classe de toutes les fonctions harmoniques [11].

c. Le problème de la membrane fournit entre autres des résultats purement mathématiques. Soit $u(r,\ t)$ une fonction harmonique à l'intérieur du cercle $\{r\leqslant 1\}$ prenant sur la circonférence $r=1$ les valeurs continues données $f(t)$. D'après ce qu'on a démontré dans *b* et compte tenu de la remarque sur l'unicité, elle s'exprime

pour $r < 1$ par l'intégrale de Poisson (12). Les coefficients de Fourier a_n et b_n de la fonction $f(t)$ tendant vers zéro pour $n \to \infty$, la série de Taylor (11) a au moins 1 pour rayon de convergence, donc représente dans le cercle $r < 1$ une fonction analytique à laquelle la fonction (12) (la fonction harmonique donnée) sert de partie réelle. Quant à la partie imaginaire de la série (11), elle nous donne une fonction harmonique $v(r, t)$ conjuguée de $u(r, t)$. Ainsi, *toute fonction harmonique admet une fonction harmonique conjuguée*. La forme explicite de la fonction $v(r, t)$ s'obtient en remplaçant dans la formule (12) le noyau de Poisson $P_r(t)$ par sa fonction harmonique conjuguée. Cette dernière est la somme des fonctions conjuguées des termes de la série (10), notamment

$$\frac{1}{2\pi} \sum_{1}^{\infty} r^n \sin nt = \frac{1}{2\pi} \frac{r \sin t}{1 - 2r \cos t + r^2}$$

(6.47*f*). Définitivement, pour la fonction cherchée $v(r, t)$, nous avons la formule

$$v(r, t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\tau) \frac{r \sin \tau}{1 - 2r \cos (t - \tau) + r^2} d\tau.$$

§ 14.5. Divergence des séries de Fourier et sommation généralisée

14.51. Si $f(t)$ est une fonction continue, alors le problème de la convergence des sommes symétriques de sa série de Fourier, sans supposition que la condition de Dini soit vérifiée, reste pour le moment ouvert. Il s'avère qu'il existe des fonctions continues dont les sommes symétriques de la série de Fourier divergent (au moins en des points isolés). Cela résulte, en fin de compte, du fait que les noyaux de Dirichlet ne forment pas une suite en forme de delta ou, plus précisément, du fait que

$$\sup_n \int_{-\pi}^{\pi} |D_n(t)| \, dt = \infty,$$

ce que nous allons voir.

Nous savons qu'une somme partielle symétrique $s_n(t)$ de la série de Fourier d'une fonction $f(t)$ peut être écrite comme suit (14.37(1))

$$s_n(f, t) = \int_{-\pi}^{\pi} f(t + h) D_n(h) \, dh,$$

où

$$D_n(h) = \frac{1}{2\pi} \frac{\sin\left(n+\frac{1}{2}\right)h}{\sin\frac{h}{2}}.$$

Posons pour simplifier $t = 0$ de sorte que

$$s_n(f, 0) = \int_{-\pi}^{\pi} f(h)\, D_n(h)\, dh.$$

C'est une suite de fonctionnelles linéaires sur l'espace de Banach $C^*(Q)$ de toutes les fonctions complexes continues sur $[-\pi, \pi]$. Nous allons voir que *les normes de ces fonctionnelles*, i.e. les nombres

$$\int_{-\pi}^{\pi} |D_n(h)|\, dh$$

(12.71*m*), *tendent vers l'infini*. Il en résultera, d'après le théorème de Banach-Steinhaus (12.74*a*), qu'il existe un élément de l'espace $C^*(Q)$, i.e. une fonction continue $f_0(t)$, pour lequel les nombres $s_n(f_0, 0)$ ne sont pas bornés. Or, cela veut dire que la série de Fourier de la fonction $f_0(t)$ n'est pas convergente (même symétriquement) au point $t = 0$.

Donc, le problème se réduit à l'établissement de la relation

$$\sup_n \int_{-\pi}^{\pi} |D_n(h)|\, dh = \infty.$$

En appliquant l'inégalité $\sin\frac{h}{2} \leqslant \frac{h}{2}$ $(0 \leqslant h \leqslant \pi)$ nous écrivons

$$\int_{-\pi}^{\pi} |D_n(h)|\, dh = \frac{1}{\pi} \int_0^{\pi} \frac{\left|\sin\left(n+\frac{1}{2}\right)h\right|}{\sin\frac{h}{2}}\, dh \geqslant$$

$$\geqslant \frac{2}{\pi} \int_0^{\pi} \frac{\left|\sin\left(n+\frac{1}{2}\right)h\right|}{h}\, dh.$$

Effectuons la substitution $(n + 1/2)\, h = t$; nous aurons

$$\int_{-\pi}^{\pi} |D_n(h)|\, dh \geqslant \frac{2}{\pi} \int_0^{(n+1/2)\pi} \frac{|\sin t|}{t}\, dt.$$

La dernière quantité croît indéfiniment pour $n \to \infty$ en vertu de la divergence de l'intégrale impropre de $|\sin t|/t$ sur l'intervalle $(0, \infty)$ (11.16*a*).

Cela justifie notre construction *).

Soulignons qu'une fonction $f_0(t)$ avec la série de Fourier divergente existe dans toute boule $U_\rho(g) = \{f : \| f - g \| \leqslant \rho\}$ de l'espace $C^*(Q)$. Pour toute fonction de ce genre, la série des modules des coefficients de Fourier diverge elle aussi, puisque la convergence de cette série impliquerait, selon le critère de Weierstrass, la convergence uniforme de la série de Fourier.

14.52. Une question se pose: ne peut-on pas y remédier en utilisant quelques méthodes de sommation de séries divergentes (12.76)? Considérons la méthode des moyennes arithmétiques (Cesàro) qui consiste à passer de la suite initiale $s_1, s_2, \ldots, s_n, \ldots$ à la suite

$$\sigma_n = \frac{s_1 + \ldots + s_n}{n} \quad (n = 1, 2, \ldots).$$

Ici le résultat positif vient d'emblée:

Théorème (L. Fejér, 1905). *Pour toute fonction $f(t)$ continue sur la circonférence $Q = \{-\pi \leqslant t \leqslant \pi\}$, la suite des sommes partielles symétriques de la série de Fourier $s_m(t) = \sum_{k=-m}^{m} c_k e^{ikt}$ converge vers $f(t)$ uniformément sur Q au sens de Cesàro, c'est-à-dire que l'on a*

$$\text{C-lim}_{m\to\infty} s_m(t) = \lim_{n\to\infty} \frac{s_0(t) + s_1(t) + \ldots + s_{n-1}(t)}{n} = f(t)$$

uniformément sur Q.

Démonstration. Conformément à 14.37, on a

$$s_m(t) = \int_{-\pi}^{\pi} f(\tau)\, D_m(t - \tau)\, d\tau,$$

par conséquent,

$$\sigma_n(t) = \frac{1}{n} \sum_{m=0}^{n-1} s_m(t) = \frac{1}{n} \int_{-\pi}^{\pi} f(\tau) \sum_{m=0}^{n-1} D_m(t - \tau)\, d\tau.$$

*) En vertu du récent théorème de Carleson (1966), les points de divergence de la série de Fourier d'une fonction $f(t) \in H(-\pi, \pi)$ représentent une exception: pour tout $\varepsilon > 0$, tous ces points peuvent être recouverts par une famille dénombrable d'intervalles dont la somme des longueurs est inférieure à ε.

Ensuite on a

$$\sum_{m=0}^{n-1} D_m(h) = \frac{1}{2\pi} \sum_{m=0}^{n-1} \frac{\sin\left(m+\frac{1}{2}\right)h}{\sin\frac{h}{2}} =$$

$$= \frac{1}{2\pi} \sum_{m=0}^{n-1} \frac{\sin\left(m+\frac{1}{2}\right)h \sin\frac{h}{2}}{\sin^2\frac{h}{2}} =$$

$$= \frac{1}{2\pi} \sum_{m=0}^{n-1} \frac{\cos mh - \cos(m+1)h}{2\sin^2\frac{h}{2}} = \frac{1}{2\pi} \frac{1-\cos nh}{2\sin^2\frac{h}{2}} = \frac{1}{2\pi} \frac{\sin^2\frac{n}{2}h}{\sin\frac{h}{2}},$$

Donc,

$$\sigma_n(t) = \frac{1}{2\pi n} \int_{-\pi}^{\pi} f(\tau) \frac{\sin^2 n\frac{t-\tau}{2}}{\sin^2\frac{t-\tau}{2}} d\tau. \tag{1}$$

La fonction

$$F_n(h) = \frac{1}{2\pi n} \frac{\sin^2\frac{n}{2}h}{\sin^2\frac{h}{2}}$$

s'appelle *noyau de Fejér*. Contrairement au noyau de Dirichlet, cette fonction est non négative. Ensuite, si $f(t) \equiv 1$, alors $s_n(t) \equiv 1$ et $\sigma_n(t) \equiv 1$, et il résulte de (1) que

$$\int_{-\pi}^{\pi} F_n(h)\, dh = 1.$$

Enfin, pour tout $\delta > 0$, on a

$$\int_{|h|>\delta} F_n(t-\tau)\, d\tau = \int_{|h|>\delta} F_n(h)\, dh \leqslant$$

$$\leqslant \frac{1}{2\pi n} \int_{|h|>\delta} \frac{dh}{\sin^2\frac{h}{2}} = \frac{C(\delta)}{2\pi n} \to 0.$$

Ainsi, le noyau de Fejér possède toutes les propriétés d'une suite en forme de delta (12.55*a*). En appliquant le théorème fondamental 12.55*d* sur les suites en forme de delta nous voyons que $\sigma_n(t) \to f(t)$ uniformément par rapport à $t \in Q$, ce qu'il fallait démontrer.

14.53. La sommation par la méthode des moyennes arithmétiques est un cas particulier de la sommation à l'aide d'une matrice de Toeplitz (12.76*c*). On se demande naturellement quelles sont les

conditions à imposer à une matrice de Toeplitz $T = \| q_{nm} \|$ pour qu'elle somme la série de Fourier de toute fonction continue $f(t)$ vers cette fonction elle-même. Rappelons que nous avons appliqué plus haut les matrices de Toeplitz à la sommation de suites bornées; or, une suite quelconque de sommes partielles de la série de Fourier d'une fonction continue n'est pas forcément bornée (14.51); nous ne considérerons donc que les matrices triangulaires de Toeplitz dont la n-ième ligne contient au plus n éléments non nuls et qui permettent, par conséquent, de construire, d'après *n'importe quelle* suite numérique $c = \{c_0, c_1, c_2, \ldots\}$, les fonctionnelles

$$T_n(c) = \sum_{m=0}^{\infty} q_{nm} c_m = \sum_{m=0}^{n} q_{nm} c_m.$$

Par définition, T-lim $c_n = \lim T_n(c)$ si la dernière limite existe.

Soient donc $T = \| q_{nm} \|$ une matrice triangulaire de Toeplitz et $Q = \sup_n \sum_{m=1}^{\infty} |q_{nm}|$. Soit

$$s(t) = \{s_m(t)\}, \quad s_m(t) = \int_{-\pi}^{\pi} f(\tau) D_m(t-\tau)\, d\tau \qquad (m = 1, 2, \ldots)$$

la suite des sommes partielles symétriques de la série de Fourier d'une fonction $f(t)$; $D_m(t)$ désigne toujours le noyau de Dirichlet

$$D_m(t) = \frac{1}{2\pi} \frac{\sin\left(m + \frac{1}{2}\right) t}{\sin \frac{t}{2}}.$$

Nous avons

$$T_n(s) = \sum_{m=0}^{n} q_{nm} s_m(t) = \sum_{m=1}^{n} q_{nm} \int_{-\pi}^{\pi} f(\tau) D_m(t-\tau)\, d\tau =$$

$$= \int_{-\pi}^{\pi} f(t) \left\{ \sum_{m=1}^{n} q_{nm} D_m(t-\tau) \right\} d\tau = \int_{-\pi}^{\pi} f(\tau) Q_n(t-\tau)\, d\tau,$$

où

$$Q_n(t) = \sum_{m=1}^{n} q_{nm} D_m(t).$$

Théorème (S. Nikolski, 1948). *S'il existe une constante $C > 0$ telle que, pour tout $n = 0, 1, 2, \ldots$, on a*

$$\int_{-\pi}^{\pi} |Q_n(t)|\, dt < C, \tag{1}$$

alors $\lim T_n(s(t)) = f(t)$ *uniformément par rapport à* $t \in [-\pi, \pi]$, *quelle que soit une fonction continue* $f(t)$. *Dans le cas contraire, il existe une fonction continue* $f(t)$ *pour laquelle les quantités* $T_n(s(t))$ *n'ont pas de limite, par exemple, au point* $t = 0$.

Démonstration. La condition (1) étant satisfaite, montrons que les noyaux $Q_n(t)$ constituent une suite en forme de delta.

Nous avons tout d'abord

$$\int_{-\pi}^{\pi} Q_n(t)\,dt = \sum_{m=1}^{n} q_{nm} \int_{-\pi}^{\pi} D_m(t)\,dt = \sum_{m=1}^{n} q_{nm} \to 1 \qquad (n \to \infty)$$

en vertu des propriétés du noyau de Dirichlet et des éléments de la matrice de Toeplitz. Ensuite, pour tout $\delta > 0$, nous avons

$$\left| \int_{|t| \geqslant \delta} Q_n(t)\,dt \right| = \left| \sum_{m=1}^{n} q_{nm} \int_{|t| \geqslant \delta} D_m(t)\,dt \right| \leqslant \sum_{m=1}^{n} |q_{nm}|\, D_{m\delta},$$

où

$$D_{m\delta} = \left| \int_{|t| \geqslant \delta} D_m(t)\,dt \right|.$$

Pour un δ fixe, la quantité $D_{m\delta}$ tend vers zéro lorsque $m \to \infty$ (14.32), donc est bornée; posons $D_\delta = \sup_m D_{m\delta}$. Pour un $\varepsilon > 0$ donné, trouvons un numéro m_0 de façon à avoir $D_{m\delta} < \varepsilon/(2Q)$ lorsque $m \geqslant m_0$. Choisissons ensuite un nombre N tel que $|q_{nm}| \leqslant \varepsilon/(2m_0 D_\delta)$ pour $m \leqslant m_0$, $n > N$; c'est possible en vertu des propriétés des éléments de la matrice de Toeplitz. Alors, pour $n > N$, on a

$$\left| \int_{|t| \geqslant \delta} Q_n(t)\,dt \right| \leqslant \sum_{m=0}^{m_0} |q_{nm}|\, D_{m\delta} + \sum_{m_0+1}^{n} |q_{nm}|\, D_{m\delta} \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Il en résulte que

$$\lim_{n \to \infty} \int_{|t| \geqslant \delta} Q_n(t)\,dt = 0.$$

Etant donné (1), nous voyons que $Q_n(t)$ est bien une suite en forme de delta. En appliquant le théorème fondamental sur les suites en forme de delta 12.55*d* nous aboutissons à la convergence uniforme de la suite $T_n(s(t))$ vers $f(t)$, et la première partie du théorème est démontrée. La deuxième partie se déduit du théorème de Banach-Steinhaus de la même façon que dans 14.51.

Si le noyau $Q_n(t)$ est non négatif, alors la condition (1) est remplie. Ceci résulte de la première formule de notre démonstration. En particulier, le noyau de Fejér (14.52) est justement de ce type.

14.54. En fait, nous avons appliqué dans 14.49*b* encore une méthode de sommation généralisée. A savoir, nous avons vu que si

$$\frac{a_0}{2} + \sum_1^\infty (a_n \cos nt + b_n \sin nt)$$

est la série de Fourier d'une fonction continue $f(t)$, alors la formule suivante a lieu :

$$f(t) = \lim_{r \to 1} \left\{ \frac{a_0}{2} + \sum_1^\infty (a_n \cos nt + b_n \sin nt) r^n \right\}.$$

Le deuxième membre de cette égalité peut être considéré comme un procédé de sommation généralisée de la série de Fourier. Ce procédé, appelé *sommation généralisée au sens de Poisson*, s'applique également à des fonctions $f(t)$ discontinues, bien que cela conduise à des résultats moins finis.

§ 14.6. Exemples de systèmes orthogonaux

14.61. Orthogonalisation. Le système de fonctions trigonométriques représente un exemple relativement rare d'un système orthogonal tout fait. Dans plusieurs cas, on construit des systèmes orthogonaux à partir des systèmes non orthogonaux moyennant le « procédé d'orthogonalisation » décrit dans 12.43*g*. Rappelons-le brièvement. Soit $f_1, f_2, \ldots, f_n, \ldots$ un système de vecteurs, dans un espace hilbertien H (réel ou complexe), fini ou non et linéairement indépendant en ce sens que tout sous-système fini $f_1, \ldots, f_n$ est linéairement indépendant au sens algébrique ordinaire. On détermine les vecteurs $g_1, \ldots, g_n, \ldots$ à l'aide des relations suivantes :

$$\left.\begin{aligned} g_1 &= f_1, \\ g_2 &= a_{21} f_1 + f_2, \\ g_3 &= a_{31} f_1 + a_{32} f_2 + f_3, \\ &\cdots\cdots\cdots\cdots\cdots \\ g_n &= a_{n1} f_1 + a_{n2} f_2 + a_{n3} f_3 + \ldots + f_n, \\ &\cdots\cdots\cdots\cdots\cdots \end{aligned}\right\} \qquad (1)$$

On démontre que *les constantes a_{jk} dans les formules* (1) *peuvent être choisies, et cela d'une façon unique, de sorte que les vecteurs $g_1, g_2, \ldots$ soient deux à deux orthogonaux.*

14.62. Polynômes de Legendre. Considérons dans l'espace hilbertien H $(-1, 1)$ le système des fonctions $f_0 \equiv 1$, $f_1 \equiv t, \ldots, f_n \equiv t^n, \ldots$ et appliquons-y le théorème d'orthogonalisation. Comme les fonctions $1, t, \ldots, t^n, \ldots$ sont linéairement indépendantes, l'hypothèse du théorème est remplie. Le sous-espace

$L_n = L(1, t, \ldots, t^n)$ est l'ensemble de tous les polynômes de degré $k \leqslant n$. La fonction $g_n(t)$ est un polynôme de degré n dont le coefficient principal vaut 1. Il s'avère que la formule explicite pour le polynôme $g_n(t)$ est

$$g_n(t) = C_n [(t^2-1)^n]^{(n)}, \qquad (1)$$

quel que soit n.

Pour le prouver, il suffit, vu l'unicité mentionnée dans 14.61, de constater que le polynôme (1) qui est évidemment de degré n est orthogonal aux fonctions $1, t, \ldots, t^{n-1}$.

Lemme. *La fonction $(t^2-1)^n$ est nulle aux points* 1 et -1 *de même que ses dérivées successives y compris la* $(n-1)$-ième; *sa dérivée d'ordre n est non nulle en ces points.*

La démonstration résulte de la représentation $(t^2-1)^n = (t+1)^n (t-1)^n$ et de la formule de Leibniz 8.12(3). En particulier,

$$[(t^2-1)^n]^{(n)}|_{t=1} = [(t+1)^n (t-1)^n]^{(n)}|_{t=1} = 2^n \cdot n!. \qquad (2)$$

Théorème. *Le polynôme* (1) *est orthogonal dans l'espace* H $(-1, 1)$ *aux fonctions* $1, t, \ldots, t^{n-1}$.

Démonstration. En intégrant par parties on trouve, pour $k < n$:

$$(t^k, [(t^2-1)^n]^{(n)}) = \int_{-1}^{1} t^k [(t^2-1)^n]^{(n)}\, dt =$$

$$= t^k [(t^2-1)^n]^{(n-1)}\Big|_{-1}^{+1} - k \int_{-1}^{1} t^{k-1} [(t^2-1)^n]^{(n-1)}\, dt.$$

Le terme sans intégrale s'annule d'après le lemme. On intègre par parties l'intégrale qui reste et l'on continue jusqu'à ce que l'exposant de t ne devienne nul :

$$(t^k, [(t^2-1)^n]^{(n)}) = -kt^{k-1} [(t^2-1)^n]^{(n-1)}\Big|_{-1}^{+1} +$$

$$+ k(k-1) \int_{-1}^{1} t^{k-2} [(t^2-1)^n]^{(n-2)}\, dt = \ldots$$

$$\ldots = \pm k! \int_{-1}^{1} [(t^2-1)^n]^{(n-k)}\, dt = \pm k! [(t^2-1)^n]^{(n-k-1)}\Big|_{-1}^{+1} = 0,$$

ce qu'il fallait démontrer.

Il est commode pour les calculs de remplacer les fonctions orthogonales obtenues par les fonctions proportionnelles qui valent 1 pour $t = 1$. A cette fin, posons dans (1) $C_n = 1/(2^n n!)$ en tenant

compte de (2). Les polynômes que l'on obtient ont la forme

$$P_n(t) = \frac{1}{2^n n!}[(t^2-1)^n]^{(n)} \qquad (n=0, 1, 2, \ldots) \tag{3}$$

et s'appellent *polynômes de Legendre.*

En particulier,

$$P_0(t) = 1,\ P_1(t) = t,\ P_2(t) = \frac{3}{2}\left(t^2 - \frac{1}{3}\right),\ P_3(t) = \frac{5}{2}\left(t^3 - \frac{3}{5}t\right), \ldots.$$

14.63. Trouvons la norme du polynôme de Legendre $P_n(t)$. Nous avons

$$(P_n, P_n) = \frac{1}{2^{2n}(n!)^2}\int_{-1}^{1}[(t^2-1)^n]^{(n)}[(t^2-1)^n]^{(n)}\,dt =$$

$$= \frac{1}{2^{2n}(n!)^2}[(t^2-1)^n]^{(n)}[(t^2-1)^n]^{(n-1)}\Big|_{-1}^{1} -$$

$$- \frac{1}{2^{2n}(n!)^2}\int_{-1}^{1}[(t^2-1)^n]^{(n+1)}[(t^2-1)^n]^{(n-1)}\,dt.$$

Le terme sans intégrale s'annule d'après le lemme. En continuant l'intégration par parties jusqu'à ce que l'ordre de la dérivée dans le deuxième facteur sous le signe somme ne devienne nul, nous obtenons

$$(P_n, P_n) = \frac{(-1)^n}{2^{2n}(n!)^2}\int_{-1}^{1}[(t^2-1)^n]^{(2n)}(t^2-1)^n\,dt =$$

$$= \frac{(-1)^n(2n)!}{2^{2n}(n!)^2}\int_{-1}^{1}(t-1)^n(t+1)^n\,dt.$$

Intégrons encore une fois par parties en abaissant l'exposant de $t-1$:

$$(P_n, P_n) =$$

$$= \frac{(-1)^n(2n)!}{2^{2n}(n!)^2}\left[(t-1)^n\frac{(t+1)^{n+1}}{n+1}\Big|_{-1}^{1} - n\int_{-1}^{1}(t-1)^{n-1}\frac{(t+1)^{n+1}}{n+1}\,dt\right] =$$

. . . . . .

$$= \frac{(-1)^n(2n)!(-1)^n n!}{2^{2n}(n!)^2(n+1)\ldots 2n}\int_{-1}^{1}(t+1)^{2n}\,dt = \frac{1}{2^{2n}}\frac{(t+1)^{2n+1}}{2n+1}\Big|_{-1}^{1} = \frac{2}{2n+1}.$$

Définitivement, nous avons

$$\|P_n\| = \sqrt{(P_n, P_n)} = \sqrt{\frac{2}{2n+1}}. \tag{1}$$

14.64. Développements suivant les polynômes de Legendre. A toute fonction $f(t) \in H(-1, 1)$ on peut faire correspondre sa *série de Fourier-Legendre*:

$$f(t) \sim \sum_{n=0}^{\infty} \gamma_n P_n(t). \tag{1}$$

Conformément à 14.17, les coefficients γ_n (*coefficients de Fourier-Legendre*) sont calculés d'après les formules

$$\gamma_n = \frac{(f, P_n)}{(P_n, P_n)} = \frac{2n+1}{2} \int_{-1}^{1} f(t) P_n(t)\, dt.$$

Tout comme dans 14.24 pour les séries de Fourier classiques, on démontre que la série de Fourier-Legendre (1) converge vers $f(t)$ au sens de la moyenne quadratique:

$$\left\| f(t) - \sum_{k=0}^{n} \gamma_k P_k(t) \right\|^2 = \int_{-1}^{1} \left| f(t) - \sum_{k=0}^{n} \gamma_k P_k(t) \right|^2 dt \to 0 \quad (n \to \infty).$$

On a l'égalité de Parseval

$$\| f \|^2 = \int_{-1}^{1} |f(t)|^2\, dt = \sum_{0}^{\infty} \frac{2}{2n+1} |\gamma_n|^2.$$

14,65*. Enonçons les théorèmes sur la convergence de la série de Fourier-Legendre aux points isolés et sur sa convergence uniforme analogues à ceux de § 14.3.

Une somme partielle de la série de Fourier-Legendre se met sous la forme

$$s_n(t) \equiv \sum_{0}^{n} \gamma_k P_k(t) = \sum_{0}^{n} \frac{(2k+1) P_k(t)}{2} \int_{-1}^{1} f(\tau) P_k(\tau)\, d\tau =$$

$$= \int_{-1}^{1} f(\tau) \sum_{0}^{n} (2k+1) \frac{P_k(\tau) P_k(t)}{2}\, d\tau.$$

La fonction

$$L_n(t, \tau) = \sum_{0}^{n} \frac{P_k(\tau) P_k(t)}{2} (2k+1)$$

s'appelle *noyau de Fourier-Legendre*. Il s'avère que la sommation peut être effectuée explicitement; le résultat s'appelle *identité de Christoffel-Darboux* (cf. exercice 11)

$$L_n(t, \tau) = \frac{n+1}{2} \frac{P_{n+1}(\tau) P_n(t) - P_n(\tau) P_{n+1}(t)}{t - \tau}.$$

En opérant avec le noyau de Fourier-Legendre de même qu'on l'a fait avec celui de Dirichlet, on peut démontrer le théorème suivant : *si une fonction* $f(t) \in H(-1, 1)$ *est continue pour* $t = t_0 \in (-1, 1)$ *et possède les dérivées finies* $f'(t_0 - 0)$ *et* $f'(t_0 + 0)$, *alors la série de Fourier-Legendre* 14.64(1) *converge au point* t_0 *vers la valeur* $f(t_0)$. *Cette convergence est uniforme sur tout ensemble* $E \subset [-1 + \delta, 1 - \delta]$, *où* $f'(t_0 - 0)$ *et* $f'(t_0 + 0)$ *sont bornées. A un point de discontinuité de première espèce* (toujours avec les valeurs $f'(t_0 - 0)$ et $f'(t_0 + 0)$ finies) *la série* 14.64(1) *converge vers* $\frac{1}{2}[f(t_0 - 0) + f(t_0 + 0)]$.

On peut trouver une démonstration de ce théorème dans [24].

14.66. En tant qu'exemple d'application des polynômes de Legendre à la physique mathématique signalons le problème suivant (cf. 14.49*b*). On a à résoudre l'équation de Laplace

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

dans la boule $r^2 = x^2 + y^2 + z^2 \leqslant 1$, à condition que, sur la frontière de la boule, i.e. pour $r = 1$, la fonction u prenne les valeurs données $u = f(\theta)$ qui ne dépendent que de l'angle θ formé par le vecteur $\{x, y, z\}$ avec l'axe des z. Pour construire la solution, on procède comme suit : après avoir développé la fonction $f(\theta)$ suivant les polynômes de Legendre en $\cos\theta$:

$$f(\theta) = \sum_0^\infty \gamma_n P_n(\cos\theta),$$

on obtient la fonction cherchée $u = u(r, \theta)$ d'après la formule *)

$$u(r, \theta) = \sum_0^\infty \gamma_n r^n P_n(\cos\theta).$$

14.67. Autres systèmes orthogonaux. On rencontre en physique mathématique bien des systèmes orthogonaux. Nous en signalons les plus usités. Ils s'obtiennent tous par un même procédé : sur un intervalle $-\infty \leqslant a \leqslant x \leqslant b \leqslant +\infty$, on considère une fonction $p(x)$ non négative (« fonction de poids ») qui sert à construire l'espace fonctionnel $H_{p(x)}[a, b]$ avec

$$(f, g)_{p(x)} = \int_a^b f(x)\,\overline{g(x)}\,p(x)\,dx$$

pour produit scalaire. Ensuite, on applique aux fonctions $1, x, x^2, \ldots$ le procédé d'orthogonalisation décrit d'une façon générale dans 14.61.

a. Pour $a = -1$, $b = 1$, $p(x) \equiv 1$, on obtient évidemment les polynômes de Legendre.

*) Cf. p. ox. [26].

b. Pour $a = -1$, $b = 1$, $p(x) = 1/\sqrt{1-x^2}$, on aboutit aux *polynômes de Tchébychev*:

$$T_n(x) = \cos(n \arccos x);$$

ces polynômes se transforment en les fonctions $\cos nt$ lorsqu'on fait la substitution $x = \cos t$, l'espace $H_{p(x)}(-1, 1)$ se trouvant isomorphe à l'espace $H_t(0, \pi)$.

c. Pour $a = 0$, $b = 1$, $p(x) = x^{q-1}(1-x)^{p-q}$, on obtient les *polynômes de Jacobi* (*polynômes hypergéométriques*).

d. Pour $a = -\infty$, $b = \infty$, $p(x) = e^{-x^2}$, on a les *polynômes d'Hermite*

$$H_n(x) = C_n e^{x^2} [e^{-x^2}]^{(n)}.$$

e. Pour $a = 0$, $b = \infty$, $p(x) = e^{-x}$, on a les *polynômes de Laguerre*

$$L_n(x) = C_n e^x [x^n e^{-x}]^{(n)}.$$

L a physiqus mathématique, en particulier les problèmes d'oscillations, a également recours à de nombreux systèmes orthogonaux de fonctions transcendentes (cf. [22]). La théorie générale des systèmes orthogonaux de fonctions est exposée dans [25] et [19].

Exercices

1. En développant en la série de Fourier la fonction impaire qui vaut $\pi/4$ pour $0 < x < \pi$, obtenir les égalités d'Euler:

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots = \frac{\pi}{4},$$

$$1 + \frac{1}{5} - \frac{1}{7} - \frac{1}{11} + \frac{1}{13} + \frac{1}{17} - \ldots = \frac{\pi}{4},$$

$$1 - \frac{1}{5} + \frac{1}{7} - \frac{1}{11} + \frac{1}{13} - \ldots = \frac{\pi}{2\sqrt{3}}.$$

2. En développant en la série de Fourier la fonction paire qui vaut x pour $0 < x < \pi$, obtenir les égalités d'Euler.

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots = \frac{\pi^2}{6},$$

$$1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \ldots = \frac{\pi^2}{12}.$$

3. Sommer les séries de Fourier:

a) $1 + \dfrac{\cos x}{1} + \dfrac{\cos 2x}{1 \cdot 2} + \ldots + \dfrac{\cos nx}{n!} + \ldots$

b) $\dfrac{\sin x}{1} + \dfrac{\sin 2x}{1 \cdot 2} + \ldots + \dfrac{\sin nx}{n!} + \ldots$

4. Si les sommes partielles d'une série de Fourier forment, dans l'espace $C(-\pi, \pi)$, un ensemble précompact (3.93a), alors la série de Fourier converge uniformément.

5. Démontrer la convergence des sommes symétriques d'une série de Fourier en un point au voisinage duquel la fonction développée $f(x)$ est monotone.

6. (Suite.) Démontrer que les sommes symétriques de la série de Fourier d'une fonction $f(t)$ convergent uniformément sur tout intervalle intérieur à un intervalle sur lequel la fonction $f(t)$ est continue et monotone.

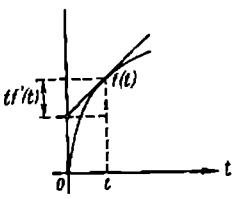


Fig. 14.7.

7. Supposons qu'une fonction $f(t)$ vérifie les conditions suivantes :

1) $f(-t) = f(t)$, $f(0) = 0$, $f(\pi - t) = f(t)$. $f(t + 2\pi) = f(t)$;

2) $f(t)$ est continue ;

3) $f'(t)$ est continue et non croissante pour $0 < t \leqslant \pi/2$;

4) $\lim\limits_{t \searrow 0} t \dfrac{f'(t)}{f(t)} = 0$; autrement dit, le segment déterminé par la tangente à la courbe $y = f(t)$ sur l'axe des ordonnées est équivalent, pour $t \searrow 0$, à l'ordonnée au point de tangence (fig. 14.7).

Montrer que les coefficients b_n dans la série de Fourier de la fonction $f(t)$

$$\sum_1^\infty b_n \sin nt$$

ont la forme :

$$b_n = 0 \text{ pour } n \text{ pair}, \quad b_n = \frac{\theta_n}{n} f\left(\frac{\pi}{n}\right) + e_n \text{ pour } n \text{ impair},$$

où $\theta_n \to 4/\pi$, $\sum\limits_1^\infty |e_n| < \infty$.

8. En utilisant la solution de l'exercice 7, donner un exemple de fonction continue ayant la série de Fourier uniformément convergente sur $[-\pi, \pi]$ et telle que la série des coefficients de Fourier ne soit pas absolument convergente.

9. En utilisant la solution de l'exercice 7, donner un exemple de fonction $f(t)$ telle que sa série de Fourier

$$\sum_{-\infty}^{\infty} c_n e^{int}$$

converge uniformément sur $[-\pi, \pi]$ tandis que chacune des séries

$$\sum_0^\infty c_n e^{int}, \qquad \sum_{-\infty}^0 c_n e^{int}$$

possède des points de divergence.

10. Soient $p(x)$ une fonction de poids (14.67) et

$$Q_n(x) = \alpha_n x^n + \beta_n x^{n-1} + \gamma_{n-2}^{(n)} x^{n-2} + \ldots + \gamma_0^{(n)}$$

la suite correspondante de polynômes orthonormés. Démontrer la formule de récurrence

$$xQ_n(x) = \frac{\alpha_n}{\alpha_{n+1}} Q_{n+1}(x) + \frac{\beta_n - \beta_{n+1}}{\alpha_n} Q_n(x) + \frac{\alpha_{n-1}}{\alpha_n} Q_{n-1}(x), \qquad (1)$$

11. (Suite.) Démontrer l'identité de Christoffel-Darboux :

$$\sum_{k=0}^{n} Q_k(x)\,Q_k(t) = \frac{\alpha_n}{\alpha_{n+1}} \frac{Q_n(x)\,Q_{n+1}(t) - Q_n(t)\,Q_{n+1}(x)}{t-x}$$

12. (Suite.) Démontrer que le polynôme $Q_n(t)$ $(n > 1)$ possède n racines dans l'intervalle $[a, b]$.

Historique

Lors de la discussion sur la corde vibrante qui s'élève aux années 1750 entre Euler et d'Alembert et dont le point principal est la définition de la fonction — est-ce une expression analytique (d'Alembert) ou bien une courbe arbitrairement tracée (Euler)? — on examine entre autres une idée de D. Bernoulli selon laquelle il serait possible de représenter n'importe quelle courbe donnée sur l'intervalle $[0, 2\pi]$ par une série de sinus et cosinus. Euler et d'Alembert ont chacun leurs raisons pour nier cette possibilité, tandis que Bernoulli ne sait pas déterminer les coefficients de sa série. La question n'est tranchée qu'en 1805, lorsque Fourier propose les formules pour les « coefficients de Fourier » (14.21*a*).

La découverte de Fourier produit un effet extraordinaire et, durant tout le XIXe siècle, est considérée comme l'un des plus remarquables théorèmes de l'analyse, bien qu'elle soit obtenue par une simple intégration terme à terme d'une série trigonométrique formellement écrite et multipliée par une fonction trigonométrique donnée. Faute de définitions rigoureuses de la convergence et de l'intégrale, Fourier ne peut démontrer la convergence de la série vers la fonction développée. C'est Dirichlet qui le fait en 1829 en se basant sur les définitions rigoureuses (Cauchy, 1821) pour les fonctions monotones par morceaux. La « condition de Dini » est formulée par Dini en 1880. Le premier exemple de divergence d'une série de Fourier pour une fonction continue est donné par Du Bois-Reymond (1876). Les « polynômes de Legendre » sont introduits par Legendre en 1785 pour résoudre l'équation de Laplace en coordonnées sphériques. Cependant, la formule explicite 14.62(3) n'est trouvée par Rodrigues qu'en 1815. Le développement suivant les polynômes de Legendre est donné par Neumann (1862) pour les fonctions analytiques et par Hobson (1908) dans le cas général. Avec les travaux de Hilbert (1906-1911), il devient possible de géométriser la théorie des développements orthogonaux.

Transformation de Fourier

Refuserais-je mon repas pour la seule raison que je ne comprends pas tout à fait le processus de digestion?

Oliver Heaviside

§ 15.1. Intégrale de Fourier et son inversion

15.11. Lorsque nous voulons représenter une fonction $\varphi(x)$ périodique de période 2π comme superposition d'harmoniques, nous nous adressons à la série de Fourier

$$\varphi(x) = \sum_{-\infty}^{\infty} a_n e^{inx}, \tag{1}$$

S'il s'agit d'une fonction de période $2\pi l$, la série de Fourier correspondante prend la forme

$$\varphi(x) = \sum_{-\infty}^{\infty} a_n e^{in\frac{x}{l}}, \tag{2}$$

où les coefficients a_n sont déterminés par la formule

$$a_n = \frac{1}{2\pi l} \int_{-\pi l}^{\pi l} \varphi(\xi)\, e^{-in\frac{\xi}{l}}\, d\xi. \tag{3}$$

La formule (3) s'obtient en multipliant (2) par $e^{-in\frac{x}{l}}$ et en intégrant ensuite par rapport à x de $-\pi l$ à πl.

Il résulte de (2) et (3) que

$$\varphi(x) = \frac{1}{2\pi} \sum_{-\infty}^{\infty} \frac{1}{l} \int_{-\pi l}^{\pi l} \varphi(\xi)\, e^{\frac{in}{l}(x-\xi)}\, d\xi. \tag{4}$$

Il est naturel d'essayer d'effectuer dans la formule (4) le passage à la limite pour $l \to \infty$, afin de représenter, si c'est possible, toute fonction $\varphi(x)$ définie sur tout l'axe $-\infty < x < \infty$ comme superposition d'harmoniques. Le passage formel à la limite amène à la

formule:

$$\varphi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\sigma \left\{ \int_{-\infty}^{\infty} \varphi(\xi)\, e^{i\sigma(x-\xi)}\, d\xi \right\}, \tag{5}$$

où le symbole σ désigne la variable continue que l'on obtient à la place de la variable discrète $\sigma_n = n/l$. Donc, la formule cherchée pour le développement d'une fonction $\varphi(x)$ suivant les harmoniques doit avoir la forme

$$\varphi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(\sigma)\, e^{i\sigma x}\, d\sigma, \tag{6}$$

où

$$\psi(\sigma) = \int_{-\infty}^{\infty} \varphi(\xi)\, e^{-i\sigma\xi}\, d\xi. \tag{7}$$

Nous avons déjà vu la formule (7) dans le chapitre sur les intégrales impropres (11.32); rappelons que la fonction $\psi(\sigma)$ définie par la formule (7) s'appelle *transformée de Fourier* (ou *intégrale de Fourier*) de la fonction $\varphi(x)$. La formule (6) est appelée *formule d'inversion* de Fourier; on dit aussi qu'elle définit la transformation *inverse* de Fourier. La transformation inverse (6) ne diffère en fait de la transformation (7) que par le signe de l'exposant et le coefficient $1/(2\pi)$.

15.12. Au lieu d'établir la légitimité du passage à la limite dans la formule 15.11(5), nous allons montrer directement que 15.11(7) implique 15.11(5), certaines conditions étant imposées à la fonction $\varphi(x)$.

On suppose en premier lieu que *la fonction* $\varphi(x)$ *soit continue par morceaux et absolument intégrable sur tout l'axe* $-\infty < x < \infty$. Cela assure l'existence de l'intégrale 15.11(7) pour toute valeur de σ, $-\infty < \sigma < \infty$.

Voici la première conséquence de la supposition faite: *la fonction* $\psi(\sigma)$ *est bornée et continue pour tout* σ *et tend vers zéro lorsque* $|\sigma| \to \infty$. La première affirmation découle de l'estimation

$$|\psi(\sigma)| \leqslant \int_{-\infty}^{\infty} |\varphi(\xi)|\, d\xi$$

L'intégrabilité absolue de la fonction $\varphi(x)$ implique la convergence uniforme par rapport au paramètre $\sigma \in (-\infty, \infty)$ de l'intégrale de Fourier 15.11(7) d'après le critère 11.47*a*. Vu le théorème 11.43 et la continuité de la fonction $e^{i\sigma\xi}$, il en résulte la continuité de la fonction $\psi(\sigma)$. Pour démontrer la troisième affirmation, trouvons,

pour un $\varepsilon > 0$ donné, un nombre A de sorte que

$$\int_{-\infty}^{-A} |\varphi(\xi)|\,d\xi + \int_{A}^{\infty} |\varphi(\xi)|\,d\xi < \frac{\varepsilon}{2}.$$

Maintenant appliquons le lemme 14.32 à l'intervalle $[-A, A]$; nous verrons qu'il existe un σ_0 tel que, pour $|\sigma| > \sigma_0$, on a

$$\left| \int_{-A}^{A} \varphi(\xi)\, e^{-i\sigma\xi}\, d\xi \right| < \frac{\varepsilon}{2}.$$

Par conséquent, pour $|\sigma| > \sigma_0$, on a

$$\left| \int_{-\infty}^{\infty} \varphi(\xi)\, e^{-i\sigma\xi}\, d\xi \right| \leqslant \int_{-\infty}^{-A} |\varphi(\xi)|\,d\xi + \left| \int_{-A}^{A} \varphi(\xi)\, e^{-i\sigma\xi}\, d\xi \right| + \int_{A}^{\infty} |\varphi(\xi)|\,d\xi < \varepsilon,$$

ce qu'il nous fallait.

15.13. Avant de démontrer la formule 15.11(5), considérons une intégrale impropre spéciale de 3e espèce:

$$I_{pq} = \int_{-\infty}^{\infty} \frac{e^{iqt} - e^{ipt}}{t}\, dt \quad (p,\ q > 0).$$

La fonction sous le signe somme est continue sur toute la droite réelle (on trouve aisément la valeur limite $i(p+q)$ pour $t = 0$). La convergence de l'intégrale I_{pq} résulte de 11.17*b*. Calculons-la. Nous avons:

$$I_{pq} = \int_{-\infty}^{\infty} \left\{ \frac{\cos qt - \cos pt}{t} + i\,\frac{\sin qt}{t} + i\,\frac{\sin pt}{t} \right\} dt =$$

$$= \int_{-\infty}^{\infty} \frac{\cos qt - \cos pt}{t}\, dt + i \int_{-\infty}^{\infty} \frac{\sin qt}{t}\, dt +$$

$$+ i \int_{-\infty}^{\infty} \frac{\sin pt}{t}\, dt = 2\pi i. \qquad (1)$$

La première intégrale s'est annulée à cause de l'imparité de la fonction à intégrer, pour la deuxième et la troisième on a utilisé la formule 11.33(1).

D'une manière analogue, on a

$$\int_{-T}^{T} \frac{e^{iqt} - e^{-ipt}}{t}\, dt = i \int_{-T}^{T} \frac{\sin qt}{t}\, dt + i \int_{-T}^{T} \frac{\sin pt}{t}\, dt.$$

Comme l'intégrale impropre

$$\int_{-\infty}^{\infty} \frac{\sin \alpha t}{t}\, dt$$

converge uniformément par rapport au paramètre $\alpha \geqslant \alpha_0 > 0$ (11.49*b*), on peut trouver pour tout $\varepsilon > 0$ un T_0 tel que, quels que soient $p \geqslant 1$, $q \geqslant 1$ et $T \geqslant T_0$, on a

$$\left| \int_{|t| \geqslant T} \frac{e^{iqt} - e^{-ipt}}{t}\, dt \right| < \varepsilon. \tag{2}$$

15.14. En passant à la démonstration de la formule 15.11(5) formulons rigoureusement la proposition à démontrer :

Théorème. *Soit $\varphi(x)$ une fonction continue par morceaux, absolument intégrable sur la droite $-\infty < x < \infty$ et vérifiant, pour un x, la condition de Dini : il existe un $\delta > 0$ tel que*

$$\int_{|t| \leqslant \delta} \left| \frac{\varphi(x+t) - \varphi(x)}{t} \right| dt < \infty.$$

Alors on a

$$\varphi(x) = \lim_{\substack{p \to \infty \\ q \to \infty}} \frac{1}{2\pi} \int_{\sigma=-p}^{q} \left\{ \int_{-\infty}^{\infty} \varphi(\xi)\, e^{i\sigma(x-\xi)}\, d\xi \right\} d\sigma, \tag{1}$$

de sorte que la limite dans le deuxième membre existe lorsque p et q tendent vers l'infini indépendamment l'un de l'autre.

Démonstration. Pour $p > 0$, $q > 0$ quelconques, nous posons

$$\varphi_{1;\, p, q}(x) = \frac{1}{2\pi} \int_{\sigma=-p}^{q} \left\{ \int_{-\infty}^{\infty} \varphi(\xi)\, e^{i\sigma(x-\xi)}\, d\xi \right\} d\sigma.$$

L'intégrale entre accolades converge uniformément par rapport à σ, et l'on peut intervertir l'ordre d'intégrations en σ et ξ d'après 11.44 :

$$\varphi_{p, q}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(\xi) \left\{ \int_{-p}^{q} e^{i\sigma(x-\xi)}\, d\sigma \right\} d\xi =$$

$$= \frac{1}{2\pi i} \int_{-\infty}^{\infty} \varphi(\xi)\, \frac{e^{iq(x-\xi)} - e^{-ip(x-\xi)}}{x-\xi}\, d\xi = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \varphi(x+t)\, \frac{e^{iqt} - e^{-ipt}}{t}\, dt.$$

La dernière transformation s'effectue à l'aide de la substitution $x - \xi = t$. Avec 15.13(1), la différence $\varphi_{p, q}(x) - \varphi(x)$ peut être

notée comme suit :

$$\varphi_{p,q}(x) - \varphi(x) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} |\varphi(x+t) - \varphi(x)| \frac{e^{iqt} - e^{-ipt}}{t}\, dt. \qquad (2)$$

Partageons cette intégrale en deux :

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} = \frac{1}{2\pi i} \int_{|t| \leqslant T} + \frac{1}{2\pi i} \int_{|t| \geqslant T} \qquad (T \gg 1). \qquad (3)$$

Le deuxième terme se met sous la forme

$$\frac{1}{2\pi i} \int_{|t| \geqslant T} \varphi(x+t) \frac{e^{iqt} - e^{-ipt}}{t}\, dt - \frac{\varphi(x)}{2\pi i} \int_{|t| \geqslant T} \frac{e^{iqt} - e^{-ipt}}{t}\, dt. \qquad (4)$$

Comme $\varphi(x+t)$ est absolument intégrable en tant que fonction de t et que le facteur $\frac{e^{iqt} - e^{ipt}}{t}$ ne dépasse pas en module le nombre 2 pour $|t| \geqslant T \geqslant 1$, nous voyons que le premier terme de la différence (4) tend vers zéro, lorsque $T \to \infty$, indépendamment des valeurs de p et q supérieures, par exemple, à 1. Le deuxième terme de (4) possède la même propriété en vertu de 15.13(2).

Mettons le premier terme dans (3) sous la forme

$$\frac{1}{2\pi i} \int_{|t| \leqslant T} \frac{\varphi(x+t) - \varphi(x)}{t} (e^{iqt} - e^{ipt})\, dt.$$

La fonction $\frac{\varphi(x+t) - \varphi(x)}{t}$ étant absolument intégrable sur l'intervalle $|t| \leqslant T$ (condition de Dini !), ce terme tend vers zéro, d'après le lemme 14.32, lorsque p et q croissent, d'où

$$\lim_{\substack{p \to \infty \\ q \to \infty}} \varphi_{p,q}(x) = \varphi(x), \qquad (5)$$

ce qu'il fallait démontrer.

La démonstration réalisée établit également la convergence uniforme de l'intégrale (1) par rapport au paramètre x parcourant un ensemble borné E de la droite $-\infty < x < \infty$ si la condition de Dini a lieu uniformément sur cet ensemble (14.35) ; on le prouve de même que le théorème analogue pour les séries de Fourier.

15.15. Si la condition de Dini n'est pas satisfaite en un point x_0, alors le théorème 15.14 n'est plus valable, et l'intégrale de Fourier pour la fonction $\varphi(x)$ peut se trouver divergente. De même qu'en cas de série de Fourier, la relation

$$\varphi(x_0) = \lim \varphi_{p,q}(x_0) = \lim_{\substack{p \to \infty \\ q \to \infty}} \frac{1}{2\pi} \int_{-p}^{q} \left\{ \int_{-\infty}^{\infty} \varphi(\xi) e^{i\sigma(x-\xi)}\, d\xi \right\} d\sigma$$

ne peut être remplie qu'au sens d'une limite généralisée. Considérons d'abord l'« intégrale partielle » symétrique

$$\varphi_{p,\,p}(x)=\frac{1}{2\pi}\int_{-p}^{p}\left\{\int_{-\infty}^{\infty}\varphi(\xi)\,e^{i\sigma(x-\xi)}\,d\xi\right\}d\sigma.$$

Dans la suite, nous la désignerons par $\varphi_p(x)$. En partant de 15.14(2) nous avons pour $\varphi_p(x)$ la représentation suivante:

$$\varphi_p(x)=\frac{1}{2\pi i}\int_{-\infty}^{\infty}\varphi(x+t)\frac{e^{ipt}-e^{-ipt}}{t}\,dt=\frac{1}{\pi}\int_{-\infty}^{\infty}\varphi(x+t)\frac{\sin pt}{t}\,dt. \qquad (1)$$

Un théorème dont la démonstration est complètement analogue à celle du théorème 14.38 a lieu:

Théorème. *Soit x_0 un point de discontinuité de* 1-ère *espèce de la fonction $\varphi(x)$, de sorte qu'il existe les limites $\varphi(x_0-0)$ et $\varphi(x_0+0)$. Si les conditions unilatérales de Dini sont satisfaites, c'est-à-dire que les intégrales*

$$\int_{x_0}^{x_0+\delta}\left|\frac{\varphi(x)-\varphi(x_0+0)}{x}\right|dx,\qquad \int_{x_0-\delta}^{x_0}\left|\frac{\varphi(x)-\varphi(x_0-0)}{x}\right|dx$$

convergent pour un $\delta>0$, alors on a

$$\varphi(x_0)=\lim_{p\to\infty}\varphi_p(x_0)=\lim_{p\to\infty}\frac{1}{2\pi}\int_{-p}^{p}\left\{\int_{-\infty}^{\infty}\varphi(\xi)\,e^{i\sigma(x-\xi)}\,d\xi\right\}d\sigma.$$

15.16. Etudions à présent les moyennes arithmétiques de l'intégrale de Fourier, de même qu'on l'a fait dans 14.52 pour la série de Fourier, sans supposer satisfaite la condition de Dini. Au lieu de la moyenne arithmétique des sommes symétriques de la série de Fourier, nous considérons, naturellement, la moyenne intégrale des intégrales symétriques $\varphi_p(x)$ (15.15):

$$\sigma_N(x)=\frac{1}{N}\int_0^N\varphi_\nu(x)\,d\nu. \qquad (1)$$

En substituant la valeur de $\varphi_\nu(x)$ de 15.15(1) nous trouvons:

$$\sigma_N(x)=\frac{1}{\pi N}\int_0^N\left\{\int_{-\infty}^{\infty}\varphi(x+t)\frac{\sin\nu t}{t}\,dt\right\}d\nu=$$

$$=\frac{1}{\pi N}\int_{-\infty}^{\infty}\frac{\varphi(x+t)}{t}\left\{\int_0^N\sin\nu t\,d\nu\right\}dt=\frac{1}{\pi N}\int_{-\infty}^{\infty}\frac{\varphi(x+t)}{t}\frac{1-\cos Nt}{t}\,dt=$$

$$=\frac{2}{\pi N}\int_{-\infty}^{\infty}\varphi(x+t)\frac{\sin^2\frac{N}{2}t}{t^2}\,dt.$$

L'expression

$$F_N(t) = \frac{2}{\pi N} \frac{\sin^2 \frac{N}{2} t}{t^2} \tag{2}$$

s'appelle *noyau de Fejér pour l'intégrale de Fourier*. Le noyau de Fejér possède les propriétés suivantes :

1) $F_N(t) \geqslant 0$;

2) $\int\limits_{-\infty}^{\infty} F_N(t)\, dt = 1$;

3) $\int\limits_{|t| \geqslant \delta} F_N(t)\, dt \to 0$ pour $N \to \infty$ quel que soit $\delta > 0$.

L'inégalité 1) est immédiate. L'égalité 2) est établie dans 11.49*b*. La relation 3) résulte de l'estimation

$$\int\limits_{|t| \geqslant \delta} F_N(t)\, dt \leqslant \frac{2}{\pi N} \int\limits_{|t| \geqslant \delta} \frac{dt}{t^2} = \frac{4}{\pi N \delta}.$$

L'égalité 2) implique la relation

$$\sigma_N(x) - \varphi(x) = \int\limits_{-\infty}^{\infty} [\varphi(x+t) - \varphi(x)]\, F_N(t)\, dt. \tag{3}$$

Nous allons démontrer le théorème suivant :

Théorème. *Si une fonction $\varphi(x)$ est absolument intégrable et uniformément continue sur un ensemble $E \subset R_1$ *), alors les moyennes arithmétiques $\sigma_N(x)$ de son intégrale de Fourier convergent uniformément sur E vers $\varphi(x)$.*

Démonstration. Pour un $\varepsilon > 0$ donné, trouvons un $\delta > 0$ tel que $|\varphi(x+t) - \varphi(x)| < \varepsilon/2$ résulte de $|t| < \delta$, $x \in E$. En appliquant les propriétés 2) et 3) du noyau de Fejér, nous avons :

$$|\sigma_N(x) - \varphi(x)| \leqslant \int\limits_{|t| \leqslant \delta} |\varphi(x+t) - \varphi(x)|\, F_N(t)\, dt + \int\limits_{|t| \geqslant \delta} |\varphi(x+t) - \varphi(x)|\, F_N(t)\, dt \leqslant$$

$$\leqslant \frac{\varepsilon}{2} \int\limits_{-\infty}^{\infty} F_N(t)\, dt + 2 \sup_{-\infty < x < \infty} |\varphi(x)| \int\limits_{|t| \geqslant \delta} F_N(t)\, dt.$$

*) La dernière propriété signifie que, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que la relation $|t| < \delta$ implique

$$|\varphi(x+t) - \varphi(x)| < \varepsilon$$

quels que soient $x \in E$ et $t \in R_1$. Soulignons que, dans cette définition, le point $x + t$ n'appartient pas forcément à l'ensemble E.

Le premier terme ne dépasse pas $\varepsilon/2$ pour tout N, le deuxième devient $< \varepsilon/2$ pour N suffisamment grand, par exemple pour $N > N_0$. Définitivement, pour $N > N_0$, nous avons

$$| \sigma_N (x) - \varphi (x) | < \varepsilon$$

quel que soit $x \in E$, ce qui démontre le théorème.

15.17. En conséquence, nous obtenons le théorème sur l'unicité de la transformée de Fourier :

Si la transformée de Fourier $\psi (\sigma)$ d'une fonction $\varphi (x)$ absolument intégrable et continue par morceaux sur l'axe $-\infty < x < \infty$ est nulle pour tous les σ, alors $\varphi (x)$ est nulle partout, à l'exception éventuelle d'un ensemble n'ayant aucun point limite fini sur l'axe des x.

En effet, on a $\psi (\sigma) \equiv 0$, $\varphi_v (x) \equiv 0$, $\sigma_N (x) \equiv 0$, donc $\varphi (x) = \lim\limits_{N \to \infty} \sigma_N (x) \equiv 0$ à l'intérieur de tout intervalle fini de continuité de $\varphi (x)$; les points de discontinuité de la fonction $\varphi (x)$ formant un ensemble au plus fini sur chaque intervalle fini de l'axe des x, cela démontre le théorème.

§ 15.2. Autres propriétés de l'intégrale de Fourier

Dans ce qui suit, nous désignons par F l'opérateur de Fourier :

$$F [\varphi (x)] = \int_{-\infty}^{\infty} \varphi (x)\, e^{-i\sigma x}\, dx.$$

L'inversion de l'intégrale de Fourier se note F^{-1} :

$$F^{-1} [\psi (\sigma)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi (\sigma)\, e^{i\sigma x}\, d\sigma.$$

15.21. Liaison entre le comportement d'une fonction $\varphi (x)$ pour $|x| \to \infty$ et la dérivabilité de sa transformée de Fourier. Nous savons que la transformée de Fourier $\psi (\sigma)$ d'une fonction $\varphi (x)$ absolument intégrable est une fonction bornée et continue de σ, $-\infty < \sigma < \infty$, qui tend vers zéro pour $|\sigma| \to \infty$. Supposons maintenant que non seulement $\varphi (x)$, mais aussi $x\varphi (x)$ soit absolument intégrable sur l'axe $-\infty < x < \infty$. On peut alors conclure que la fonction $\psi (\sigma)$ est dérivable. En effet, la dérivation formelle par rapport au paramètre σ de l'intégrale de Fourier

$$\int_{-\infty}^{\infty} \varphi (x)\, e^{-i\sigma x}\, dx = \psi (\sigma)$$

amène à l'intégrale

$$-i\int_{-\infty}^{\infty} x\varphi(x)\,e^{-i\sigma x}\,dx$$

qui est absolument convergente et uniformément convergente par rapport à σ. En vertu du théorème 11.45*a*, la fonction $\psi(\sigma)$ est dérivable et l'on a

$$\psi'(\sigma) = -i\int_{-\infty}^{\infty} x\varphi(x)\,e^{-i\sigma x}\,dx.$$

On aboutit à la formule éloquente

$$iF'[\varphi] = F[x\varphi] \qquad (1)$$

qui montre que l'opération de multiplication par x est transformée par l'opérateur de Fourier en l'opération $i\frac{d}{d\sigma}$. En tant que transformée d'une fonction absolument intégrable, la fonction $\psi'(\sigma)$ est toujours continue, bornée et tend vers zéro pour $|\sigma| \to \infty$. Si les fonctions $x\varphi(x)$, $x^2\varphi(x)$, ..., $x^m\varphi(x)$ sont absolument intégrables sur l'axe des x avec la fonction $\varphi(x)$, on peut continuer la dérivation; on verra que la fonction $\psi(\sigma) = F[\varphi]$ possède ses dérivées successives, y compris la m-ième, continues, bornées et tendant vers zéro pour $|\sigma| \to \infty$; on a la formule

$$i^k F^{(k)}[\varphi] = F[x^k\varphi] \qquad (k = 0, 1, \ldots, m). \qquad (2)$$

Si *tous* les produits $x^m\varphi(x)$ sont absolument intégrables ($m = 0, 1, \ldots$), alors la fonction $F[\varphi] = \psi(\sigma)$ possède *toutes* les dérivées en σ, chaque dérivée étant continue et tendant vers zéro lorsque $|\sigma| \to \infty$.

Nous voyons que plus la décroissance à l'infini de la fonction $\varphi(x)$ est rapide, plus la fonction $\psi(\sigma)$ est lisse.

15.22. Voyons comment se perfectionnent les propriétés de dérivabilité de la fonction $\psi(\sigma)$ lorsqu'on impose des conditions supplémentaires au comportement de la fonction $\varphi(x)$ à l'infini.

a. Supposons que c'est le produit $\varphi(x)\,e^{b|x|}$, avec une constante fixe $b > 0$, qui soit intégrable. Alors on peut dire que *la transformée de Fourier* $\psi(\sigma)$ *de la fonction* $\varphi(x)$ *est une fonction non seulement indéfiniment dérivable, mais aussi analytique*. En effet, l'intégrale de Fourier

$$\psi(\sigma) = \int_{-\infty}^{\infty} \varphi(x)\,e^{-i\sigma x}\,dx$$

est maintenant définie non seulement pour les σ réels, mais aussi pour certains σ complexes: si l'on pose $s = \sigma + i\tau$ (σ, τ réels), on a

$$\psi(\sigma + i\tau) = \int_{-\infty}^{\infty} \varphi(x)\, e^{-i\sigma x} e^{\tau x}\, dx = \int_{-\infty}^{\infty} \varphi(x)\, e^{-isx}\, dx, \qquad (1)$$

et l'intégrale converge pour $|\tau| \leqslant b$, i.e. dans toute une bande horizontale du s-plan complexe. La fonction obtenue de la variable complexe s est analytique en tout point intérieur de la bande; en effet, l'intégrale converge uniformément dans un voisinage du point s (contenu entièrement dans la bande), ce qui permet d'appliquer le théorème 11.45*b*. La fonction $\psi(s)$ est bornée dans toute la bande car

$$|\psi(s)| \leqslant \int_{-\infty}^{\infty} |\varphi(x)|\, e^{|\tau||x|}\, dx \leqslant \int_{-\infty}^{\infty} |\varphi(x)|\, e^{b|x|}\, dx.$$

On peut dire que *la fonction* $\psi(s) = \psi(\sigma + i\tau)$ *converge vers zéro uniformément par rapport à* τ, $|\tau| \leqslant b$, *lorsque* $\sigma \to \pm\infty$.

Pour le démontrer, on doit préciser un peu les raisonnements de 15.12. Notamment, comme la fonction $\varphi(x)\, e^{b|x|}$ est absolument intégrable, on peut choisir, pour un $\varepsilon > 0$ donné, un nombre A tel que

$$\int_{-\infty}^{-A} |\varphi(x)|\, e^{b|x|}\, dx + \int_{A}^{\infty} |\varphi(x)|\, e^{bx}\, dx < \frac{\varepsilon}{2}.$$

Considérons l'intégrale

$$\int_{-A}^{A} \varphi(x)\, e^{-isx}\, dx = \int_{-A}^{A} \varphi(x)\, e^{\tau x} e^{-i\sigma x}\, dx.$$

D'après l'inégalité 14.32(5), elle admet la majoration

$$\left| \int_{-A}^{A} \varphi(x)\, e^{-isx}\, dx \right| \leqslant 2A\omega\left[\varphi(x)\, e^{\tau x}, \frac{2\pi}{|\sigma|}\right] + N_A M_A \frac{2\pi}{|\sigma|}, \qquad (2)$$

où $\omega(\psi, \delta)$ désigne l'oscillation de la fonction $\psi(x)$ sur ses intervalles de continuité, N_A le nombre de tels intervalles pour la fonction $\varphi(x)$ sur $[-A, A]$, $M_A = \sup_{|x| \leqslant A} |\varphi(x)|\, e^{\tau x}$.

Le premier terme dans le deuxième membre de (2) ne dépasse pas, pour $|\tau| \leqslant b$, la quantité (5.17*d*)

$$2A\omega\left[\varphi(x), \frac{2\pi}{|\sigma|}\right] e^{Ab} + 2A\omega\left[e^{bx}, \frac{2\pi}{|\sigma|}\right] \max_{|x| \leqslant A} |\varphi(x)|$$

qui tend vers 0, lorsque $|\sigma| \to \infty$, indépendamment de la valeur de τ, $|\tau| \leqslant b$. Il en est de même, évidemment, du second terme dans (2). On peut choisir un σ_0 de façon que, pour $|\sigma| > \sigma_0$, $|\tau| \leqslant b$,

on ait

$$\left| \int_{-A}^{A} \varphi(x)\, e^{-i\sigma x}\, dx \right| < \frac{\varepsilon}{2}.$$

Il en résulte pour $|\sigma| > \sigma_0$, tout comme dans 15.12 :

$$\left| \int_{-\infty}^{\infty} \varphi(x)\, e^{-i\sigma x}\, dx \right| < \varepsilon,$$

ce qu'il nous fallait.

b. Supposons à présent que le produit de la fonction $\varphi(x)$ par $e^{b|x|}$ soit intégrable pour *tout* b. Alors la fonction $\psi(s)$ est définie et analytique dans toute bande $|\tau| \leqslant b$, i.e. est une *fonction analytique entière* ; d'après ce qu'on a vu, cette fonction entière *reste bornée* et uniformément convergente vers zéro pour $\sigma \to \pm\infty$ *dans toute bande* $|\tau| \leqslant b$ (avec le majorant dépendant de b).

15.23. On peut considérer les fonctions $\varphi(x)$ qui décroissent à l'infini d'une façon encore plus rapide, notamment les fonctions pour lesquelles est intégrable le produit $\varphi(x)\, e^{M(x)}$, où $M(x)$ croît plus rapidement que toute fonction linéaire. Il est commode de mettre $M(x)$ sous la forme

$$M(x) = \int_0^x \mu(\xi)\, d\xi \qquad (0 \leqslant x < \infty), \tag{1}$$

où $\mu(\xi)$ est une fonction continue croissante avec $\mu(0) = 0$, $\mu(\infty) = \infty$; pour x négatif, nous posons $M(x) = M(-x)$.

Dans le présent cas, pour décrire les propriétés de la transformée de Fourier de $\varphi(x)$, on peut utiliser la fonction $\Omega(\tau)$ duale au sens de Young de la fonction $M(x)$. A savoir, on dit *duale au sens de Young* d'une fonction $M(x)$ la fonction $\Omega(\tau)$ définie par les égalités

$$\Omega(\tau) = \int_0^\tau \lambda(t)\, dt \qquad (0 \leqslant \tau < \infty),\ \Omega(-\tau) = \Omega(\tau). \tag{2}$$

où $\lambda(t)$ est la fonction inverse de $\mu(\xi)$. Les fonctions duales au sens de Young sont liées par l'inégalité de Young (9.61*h*) :

$$x\tau \leqslant M(x) + \Omega(\tau) \quad (x > 0,\ \tau > 0). \tag{3}$$

Théorème. *Si* $\varphi(x)$ *est une fonction continue par morceaux pour laquelle l'intégrale*

$$\int_{-\infty}^{\infty} |\varphi(x)|\, e^{M(x)}\, dx \tag{4}$$

est finie, alors la transformée de Fourier $\psi(s)$ *de la fonction* $\varphi(x)$ *est une fonction analytique entière vérifiant l'inégalité*

$$|\psi(\sigma+i\tau)| \leqslant Ce^{\Omega(\tau)}. \tag{5}$$

Démonstration. Le fait que $\psi(s)$ est une fonction analytique entière résulte de 15.22*b*. Ensuite, nous avons

$$|\psi(\sigma+i\tau)| = \left|\int_{-\infty}^{\infty} \varphi(x)\, e^{-i(\sigma+i\tau)x}\, dx\right| \leqslant \int_{-\infty}^{\infty} |\varphi(x)|\, e^{M(x)} e^{|\tau||x|} e^{-M(x)}\, dx.$$

En appliquant à l'exposant l'inégalité de Young (3) nous obtenons

$$|x||\tau| - M(x) \leqslant \Omega(\tau),$$

d'où

$$|\psi(\sigma+i\tau)| \leqslant e^{\Omega(\tau)} \int_{-\infty}^{\infty} |\varphi(x)|\, e^{M(x)}\, dx = Ce^{\Omega(\tau)},$$

ce qui démontre le théorème.

En choisissant, par exemple, une fonction $\varphi(x)$ qui satisfait à la condition

$$\int_{-\infty}^{\infty} |\varphi(x)|\, e^{\frac{1}{p}|x|^p}\, dx < \infty, \quad p > 1,$$

on trouve que la fonction correspondante $\psi(s)$ vérifie l'inégalité

$$|\psi(\sigma+i\tau)| \leqslant Ce^{\frac{1}{q}|\tau|^q} \left(\frac{1}{p}+\frac{1}{q}=1\right),$$

car $\frac{1}{q}x^q$ est la fonction duale au sens de Young de $\frac{1}{p}x^p$ (9.61*h*). Notons que les nombres p et q sont les deux supérieurs à 1, mais varient dans les sens opposés: lorsque p croît, q décroît, et lorsque $p \to \infty$, on a $q \to 1$.

15.24. Supposons enfin que c'est le produit de $\varphi(x)$ par *toute* fonction croissante de $|x|$ qui soit intégrable. Il est aisé de voir que ce sont les *fonctions à support borné* $\varphi(x)$ (qui s'annulent presque partout à l'extérieur d'un intervalle $|x| \leqslant a$), et elles seules, qui possèdent cette propriété. Supposons donc que $\varphi(x)$ soit nulle pour $|x| \geqslant a$. Alors la transformée de Fourier

$$\psi(s) = \int_{-a}^{a} \varphi(x)\, e^{-isx}\, dx$$

est une fonction analytique entière de s qui admet dans le plan des s la majoration suivante:

$$|\psi(\sigma+i\tau)| \leqslant \int_{-a}^{a} |\varphi(x)|\, e^{|\tau||x|}\, dx \leqslant Ce^{a|\tau|}, \tag{1}$$

avec $C = \int_{-a}^{a} |\varphi(x)|\,dx$. Une fonction analytique $\psi(s)$ vérifiant l'inégalité (1) s'appelle *fonction entière de type exponentiel fini* $\leqslant a$.

Ainsi, plus la décroissance de $\varphi(x)$ à l'infini est rapide, plus sa transformée de Fourier $\psi(\sigma)$ est « lisse ». En partant des fonctions $\psi(\sigma)$ continues nous avons passé par les fonctions à plusieurs dérivées, indéfiniment dérivables, analytiques dans une bande, dans tout le plan, et nous sommes arrivés aux fonctions analytiques de type exponentiel fini. Aucune fonction tendant vers zéro dans les deux sens sur l'axe réel ne peut être plus « lisse » (or, nous savons que chaque transformée de Fourier d'une fonction intégrable possède cette propriété); on sait qu'il n'existe aucune fonction analytique entière non nulle de type exponentiel fini qui tende vers zéro sur l'axe des abscisses et croisse dans le plan moins rapidement que $e^{a|\tau|}$ pour tout $a > 0$ (cf. exercice 24 du chapitre 10).

15.25. A présent, au lieu des conditions d'une décroissance de plus en plus rapide, nous imposerons à la fonction $\varphi(x)$ les conditions de devenir de plus en plus lisse. Vu les résultats 15.21-15.24, on est en droit de s'attendre à ce que la transformée de Fourier de la fonction $\varphi(x)$ sera alors soumise aux conditions d'une décroissance de plus en plus rapide.

Supposons qu'une fonction absolument intégrable $\varphi(x)$ soit continue et possède une dérivée continue par morceaux et absolument intégrable elle aussi sur l'axe $-\infty < x < \infty$. Il en résulte avant tout que la fonction

$$\varphi(x) = \varphi(0) + \int_0^x \varphi'(\xi)\,d\xi$$

a une limite pour $x \to \infty$; cette limite est nulle, sinon $\varphi(x)$ ne serait pas intégrable. Il en est de même du cas $x \to -\infty$. Ensuite, en intégrant par parties nous avons

$$F[\varphi'] = \int_{-\infty}^{\infty} \varphi'(x)\,e^{-ix\sigma}\,dx = \varphi(x)\,e^{-ix\sigma}\Big|_{-\infty}^{\infty} + i\sigma \int_{-\infty}^{\infty} \varphi(x)\,e^{-ix\sigma}\,dx.$$

D'après ce qui précède, le premier terme à droite est nul; nous avons l'égalité

$$F[\varphi'] = i\sigma F[\varphi].$$

En d'autres termes, *la dérivation de la fonction* $\varphi(x)$ *correspond à la multiplication de la fonction* $\psi(\sigma) = F[\varphi]$ *par* $i\sigma$. Comme $F[\varphi'(x)]$, en tant que transformée de Fourier d'une fonction intégrable, est une fonction bornée de σ (et même tendant vers zéro pour $|\sigma| \to \infty$),

nous avons la majoration suivante pour $F[\varphi(x)]$:

$$F[\varphi] = \frac{|F[\varphi'(x)]|}{|\sigma|} \leqslant \frac{c}{|\sigma|}.$$

Ainsi, dans le présent cas, non seulement la fonction $\psi(\sigma)$ tend vers zéro pour $|\sigma| \to \infty$, mais elle le fait plus rapidement que $1/|\sigma|$. Si les dérivées successives y compris la m-ième de la fonction $\varphi(x)$ sont absolument intégrables, alors en continuant le procédé nous obtenons

$$F[\varphi^{(k)}(x)] = (i\sigma)^k F[\varphi] \quad (k = 0, 1, \ldots, m). \tag{1}$$

De même que plus haut, nous avons la majoration

$$F[\varphi] = \frac{|F[\varphi^{(k)}(x)]|}{|\sigma|^k} \leqslant \frac{c}{|\sigma|^k}. \tag{2}$$

Donc, *plus la fonction $\varphi(x)$ a de dérivées intégrables, plus la décroissance à l'infini de sa transformée de Fourier vers zéro est rapide.*

En particulier, lorsque la fonction $\varphi(x)$ est assez lisse, sa transformée de Fourier est elle aussi absolument intégrable. On voit de (2) que l'existence et l'intégrabilité absolue de φ, φ' et φ'' en donnent une condition suffisante.

Si $\varphi^{(k)}(x)$ existe et est absolument intégrable pour *tout* $k = 0, 1, 2, \ldots$, alors la fonction $\psi(\sigma)$ décroît, pour $|\sigma| \to \infty$, plus rapidement que *toute* fonction $1/|\sigma|^k$.

15.26. a. Supposons maintenant que la fonction $\varphi(x)$ soit non seulement indéfiniment dérivable, mais aussi analytique dans une bande $|y| \leqslant b$ du plan de la variable complexe $z = x + iy$. Supposons en outre qu'il existe une fonction $\Phi(x)$ telle que

$$\lim_{|x| \to \infty} \Phi(x) = 0, \qquad \int_{-\infty}^{\infty} \Phi(x)\, dx < \infty \tag{1}$$

et que, pour tout y $|y| \leqslant b$, on ait

$$|\varphi(x + iy)| \leqslant \Phi(x). \tag{2}$$

Nous verrons (*c*) que la transformée de Fourier de la fonction $\varphi(x)$ est alors une fonction exponentiellement décroissante.

b. Etablissons préalablement le lemme suivant de la théorie des fonctions analytiques :

L e m m e. *Si une fonction $f(z)$ est analytique dans la bande $|y| < b$ et y satisfait à l'inégalité*

$$|f(x + iy)| \leqslant \Phi(x) \tag{3}$$

(*où* $\Phi(x)$ *est une fonction vérifiant les conditions* (1)), *alors l'intégrale*

$$\int_{-\infty}^{\infty} f(x+iy)\,dx \tag{4}$$

ne dépend pas de y, $|y| < b$.

Démonstration. L'existence de l'intégrale (4) résulte immédiatement de la continuité de la fonction $f(x+iy)$ par rapport à x et des estimations (1) et (3). Soient maintenant $y_1 < y_2$ deux

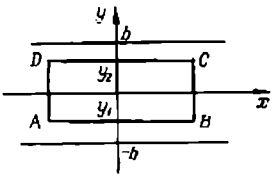


Fig. 15.1.

nombres quelconques de l'intervalle $(-b, b)$. Considérons le contour fermé $L = ABCD$ montré fig. 15.1. D'après le théorème de Cauchy 10.32 on a

$$\int_L f(z)\,dz = \int_A^B f(z)\,dz + \int_B^C f(z)\,dz + \int_C^D f(z)\,dz + \int_D^A f(z)\,dz = 0. \tag{5}$$

Soient $-R$ et R les abscisses des points A et B. Alors

$$\left|\int_B^C f(z)\,dz\right| \leqslant \int_{y_1}^{y_2} |f(z)|\,dz \leqslant \int_{y_1}^{y_2} \Phi(R)\,dy = \Phi(R)(y_2 - y_1);$$

cette quantité tend vers 0 pour $R \to \infty$, de même que l'intégrale

$$\int_D^A f(z)\,dz.$$

Les deux intégrales qui restent ont, pour $R \to \infty$, les limites respectives

$$\int_{-\infty}^{\infty} f(x+iy_1)\,dx \quad \text{et} \quad -\int_{-\infty}^{\infty} f(x+iy_2)\,dx.$$

En passant dans l'égalité (5) à la limite pour $R \to \infty$, on obtient

$$\int_{-\infty}^{\infty} f(x+iy_1)\,dx = \int_{-\infty}^{\infty} f(x+iy_2)\,dx,$$

ce qu'il fallait démontrer.

c. Théorème. *Dans l'hypothèse de a, la transformée de Fourier* $\psi(\sigma)$ *de la fonction* $\varphi(x)$ *vérifie l'inégalité*

$$|\psi(\sigma)| \leqslant Ce^{-b|\sigma|}.$$

Démonstration. Appliquons le lemme b à la fonction analytiquo

$$f(z) = \varphi(z)\, e^{-i\sigma z}$$

pour laquelle l'inégalité (3) a lieu si l'on y remplace $\Phi(x)$ par $\Phi(x)\, e^{\sigma|b|}$. En vertu du lemme, on a

$$\psi(\sigma) = \int_{-\infty}^{\infty} \varphi(x)\, e^{-i\sigma x}\, dx = \int_{-\infty}^{\infty} \varphi(x+iy)\, e^{-i\sigma(x+iy)}\, dx \qquad (6)$$

pour tout $|y| < b$. En fixant y, nous obtenons l'estimation

$$|\psi(\sigma)| \leqslant \int_{-\infty}^{\infty} |\varphi(x+iy)|\, e^{\sigma y}\, dx = e^{\sigma y} \int_{-\infty}^{\infty} |\varphi(x+iy)|\, dx \leqslant e^{\sigma y} \int_{-\infty}^{\infty} \Phi(x)\, dx.$$

En passant ici à la limite pour $y \to -b \operatorname{sgn} \sigma$, nous avons

$$|\psi(\sigma)| \leqslant Ce^{-|\sigma| b},$$

ce qu'il nous fallait.

d. Si la fonction $\varphi(x)$ est analytique dans tout le plan $z = x + iy$ et si l'on peut indiquer, pour toute bande $|y| < b$, une fonction $\Phi(x)$ vérifiant les conditions (1) et (2) (la fonction $\Phi(x)$ peut dépendre de b), alors en appliquant le théorème c nous voyons que *la transformée de Fourier* $\psi(\sigma)$ *de la fonction* $\varphi(x)$ *satisfait à une inégalité de la forme*

$$|\psi(\sigma)| \leqslant C_b e^{-|\sigma| b}$$

pour tout b.

15.27. Soit, ensuite $\varphi(x)$ une fonction analytique entière qui admet dans toute bande $|y| \leqslant b$ l'estimation

$$|\varphi(x+iy)| \leqslant e^{\Omega(y)} \Phi_b(x)$$

avec

$$\Omega(y) = \int_0^y \lambda(\eta)\, d\eta \quad (0 \leqslant y < \infty), \quad \Omega(-y) = \Omega(y),$$

où $\lambda(\eta)$ est une fonction continue croissante telle que $\lambda(0) = 0$, $\lambda(\infty) = \infty$.

Pour tout b, la fonction $\Phi_b(x)$ est supposée vérifiant les conditions 15.26(1) dans la bande $|y| \leqslant b$.

Soit $M(\sigma)$ la duale au sens de Young de la fonction $\Omega(y)$:

$$M(\sigma) = \int_0^{\sigma} \mu(\xi)\, d\xi,$$

où $\mu(\xi)$ est la fonction inverse de $\lambda(\eta)$.

Théorème. *Dans l'hypothèse formulée, la transformée de Fourier $\psi(\sigma)$ de la fonction $\varphi(x)$ satisfait à l'inégalité*

$$|\psi(\sigma)| \leqslant Ce^{-M(\sigma)}.$$

Démonstration. D'après 15.26(6), nous avons

$$\psi(\sigma) = \int_{-\infty}^{\infty} \varphi(x)\, e^{-i\sigma x}\, dx = \int_{-\infty}^{\infty} \varphi(x+iy)\, e^{-i\sigma(x+iy)}\, dx$$

pour tout y. Il en résulte l'estimation $(b > |y|)$

$$|\psi(\sigma)| \leqslant \int_{-\infty}^{\infty} e^{\Omega(y)} \Phi_b(x)\, e^{\sigma y}\, dx = C_b e^{\Omega(y)+\sigma y}. \tag{1}$$

Choisissons le signe et le module de y de façon que $\sigma y = -|\sigma|\,|y|$ et que l'inégalité de Young devienne l'égalité :

$$|\sigma|\,|y| = M(\sigma) + \Omega(y).$$

Ceci fait, il résulte de (1) que

$$|\psi(\sigma)| \leqslant C_b e^{-M(\sigma)},$$

et le théorème est démontré.

15.28. Soit, enfin, $\varphi(x)$ une fonction analytique entière qui satisfait à l'inégalité

$$|\varphi(x+iy)| \leqslant \Phi(x)\, e^{a|y|},$$

où la fonction $\Phi(x)$ est soumise aux conditions 15.26(1) (et ne dépend pas de y).

Théorème. *Dans l'hypothèse formulée, la transformée de Fourier $\psi(\sigma)$ de la fonction $\varphi(x)$ s'annule pour $|\sigma| > a$.*

Démonstration. D'après 15.26(6), nous avons

$$\psi(\sigma) = \int_{-\infty}^{\infty} \varphi(x)\, e^{-i\sigma x}\, dx = \int_{-\infty}^{\infty} \varphi(x+iy)\, e^{-i\sigma(x+iy)}\, dx \tag{1}$$

pour tout y. Fixons le signe de y de façon à avoir $\sigma y = -|\sigma|\,|y|$. Il résulte alors de (1) :

$$|\psi(\sigma)| \leqslant e^{\sigma y} \int_{-\infty}^{\infty} \Phi(x)\, e^{a|y|}\, dx \leqslant Ce^{(a-|\sigma|)|y|}. \tag{2}$$

Soit $|\sigma| > a$. En faisant tendre $|y|$ vers ∞ dans l'inégalité (2) nous avons $\psi(\sigma) = 0$, ce qu'il nous fallait.

15.29. Nous voyons que les théorèmes 15.25-15.28 ne sont pas les réciproques exactes des théorèmes 15.21-15.24 et demandent certaines hypothèses supplémentaires (par exemple, l'existence d'une fonction $\Phi(x)$ possédant les propriétés 15.26(1), (2)). La question se pose sur la construction des classes des fonctions $\varphi(x)$ dont on puisse caractériser complètement la classe des transformées de Fourier $\psi(\sigma)$. Certaines de ces classes peuvent être construites à l'aide des théorèmes 15.21-15.28.

a. Classe S. Considérons l'ensemble S de toutes les fonctions indéfiniment dérivables $\varphi(x)$ $(-\infty < x < \infty)$ dont chacune vérifie, pour n'importe quels k, $q = 0, 1, 2, \ldots$, une inégalité de la forme

$$|x^k \varphi^{(q)}(x)| \leqslant C_{kq}, \tag{1}$$

où C_{kq} est une constante (dépendant du choix de la fonction $\varphi(x)$).

Chaque fonction $x^k \varphi^{(q)}(x)$ est non seulement bornée sur l'axe des x, mais aussi intégrable sur tout l'axe car l'inégalité

$$|x^k \varphi^{(q)}(x)| \leqslant \frac{C_{k+2,\,q}}{x^2}$$

est valable avec (1), de sorte que

$$|x^k \varphi^{(q)}(x)| \leqslant \min\left\{C_{kq}, \frac{C_{k+2,\,q}}{x^2}\right\} \leqslant \frac{C^*_{kq}}{1+x^2},$$

où C^*_{kq} est une nouvelle constante.

Toute fonction $x^k \varphi(x)$ est indéfiniment dérivable avec $\varphi(x)$, et chacune de ses dérivées est intégrable sur l'axe des x, parce que s'exprimant linéairement par des fonctions intégrables $x^j \varphi^{(q-j)}(x)$ d'après la formule de Leibniz 8.12(3).

La fonction $\psi(\sigma) = F|\varphi(x)|$ est indéfiniment dérivable en vertu de 15.21. En nous servant des formules 15.21(2) et 15.25(1) nous pouvons écrire

$$F[(x^k \varphi(x))^q] = (-i)^q i^k \sigma^k \psi^{(q)}(\sigma);$$

ici le deuxième membre, en tant que transformée de Fourier d'une fonction intégrable $(x^k \varphi(x))^q$, est borné pour n'importe quels k et q:

$$|\sigma^k \psi^{(q)}(\sigma)| \leqslant B_{kq}.$$

Donc, *si* $\varphi(x) \in S$, *alors* $\psi(\sigma) \in S$. Inversement, soit $\psi(\sigma) \in S$; montrons que *cette fonction est la transformée de Fourier d'une fonction* $\varphi(x) \in S$. Posons

$$\varphi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(\sigma)\, e^{i\sigma x}\, d\sigma.$$

La fonction $2\pi\varphi(-x)$ est la transformée de Fourier de la fonction $\psi(\sigma)$, donc appartient à S. Mais alors, évidemment, on a aussi $\varphi(x) \in S$. D'après la formule d'inversion :

$$\psi(\sigma) = \frac{1}{2\pi} \int_{-\infty}^{\infty} 2\pi\varphi(-x)\, e^{i\sigma x}\, dx = \int_{-\infty}^{\infty} \varphi(x)\, e^{-i\sigma x}\, dx ;$$

par conséquent, $\psi(\sigma)$ est la transformée de Fourier de la fonction $\varphi(x)$, ce qu'il fallait démontrer.

Ainsi, *la transformation de Fourier F applique la classe S sur toute la classe S.*

b. Soit m_{kq} $(k, q = 0, 1, 2, \ldots)$ une suite double de constantes. Par définition, la classe $S_{\langle m_{kq}\rangle}$ est formée de toutes les fonctions indéfiniment dérivables $\varphi(x)$ $(-\infty < x < \infty)$ dont chacune satisfait aux inégalités

$$|x^k \varphi^{(q)}(x)| \leqslant CA^k B^q m_{kq} \quad (k, q = 0, 1, 2, \ldots),$$

où les constantes A, B et C peuvent dépendre de la fonction $\varphi(x)$.

Il s'avère que, dans certaines conditions pour la suite m_{kq}, on a la formule

$$F(S_{\langle m_{kq}\rangle}) = S_{\langle m_{kq}\rangle}\ ^{*}).$$

c. Classes W_M et W^{Ω}. Soient $M(x)$ et $\Omega(\tau)$ deux fonctions duales au sens de Young l'une de l'autre (15.23). Par définition, la classe W_M est formée de toutes les fonctions indéfiniment dérivables $\varphi(x)$ $(-\infty < x < \infty)$ dont chacune satisfait aux inégalités

$$|\varphi^{(q)}(x)| \leqslant C_q e^{-M(x)} \quad (q = 0, 1, 2, \ldots).$$

Si $\psi(x)$ est la transformée de Fourier d'une fonction $\varphi(x)$, alors $(is)^q \psi(s)$ est celle de la fonction $\varphi^{(q)}(x)$ (15.25). En vertu de 15.23, on a les inégalités

$$|s^q \psi(\sigma + i\tau)| \leqslant C'_q e^{\Omega(\tau)} \quad (q = 0, 1, 2, \ldots). \tag{2}$$

Désignons par W^{Ω} la classe de toutes les fonctions analytiques entières $\psi(s)$ vérifiant les inégalités de la forme (2). Nous voyons que $F(W_M) \subset W^{\Omega}$. Soit maintenant $\psi(s)$ une fonction quelconque de W^{Ω}. A partir de l'inégalité (2) et de la même inégalité où l'on remplace q par $q + 2$, nous obtenons

$$|s^q \psi(\sigma + i\tau)| \leqslant e^{\Omega(\tau)} \min\left\{C'_q, \frac{C'_{q+2}}{|s|^2}\right\} = \Phi_{\tau q}(\sigma)\, e^{\Omega(\tau)},$$

où

$$\Phi_{\tau q}(\sigma) = \min\left\{C'_q, \frac{C'_{q+2}}{|\sigma + i\tau|^2}\right\} \leqslant \frac{C_{\tau q}}{1 + |\sigma|^2}$$

est une fonction intégrable. L'application du théorème 15.27 amène à la conclusion que

$$|\varphi^{(q)}(x)| \leqslant C''_q e^{-M(x)},$$

*) Cf. [23].

c'est-à-dire que $F[W^{\Omega}] = W_M$. Définitivement, *l'image de la classe W_M par la transformation de Fourier est la classe W^{Ω}, et l'image de la classe W^{Ω} est W_M.*

§ 15.3. Exemples et applications

Au début, dans 15.31-15.32, nous considérons quelques exemples de transformées de Fourier.

15.31. La transformée de Fourier d'une fraction rationnelle

$$Q(x) = \frac{a_0 + a_1 x + \ldots + a_m x^m}{b_0 + b_1 x + \ldots + b_n x^n},$$

où $m < n - 1$, était calculée dans 11.32*b* par les méthodes d'intégration le long d'un contour. En utilisant le théorème 15.26*c* et l'analyticité de la fonction $Q(z)$ dans une bande $|y| \leqslant b$ (qui ne contient pas de racines du dénominateur) on pourrait affirmer que $F[Q(x)]$ décroît exponentiellement pour $|\sigma| \to \infty$; d'ailleurs, on n'en a plus besoin, $F[Q(x)]$ étant déjà calculée d'une façon explicite.

15.32. Trouvons la transformée de Fourier $\psi(\sigma)$ d'une fonction $\varphi(x) = e^{-ax^2}$, $a > 0$.

Cette fonction est analytiquement prolongeable dans tout le plan $z = x + iy$ avec la majoration

$$|e^{-az^2}| = |e^{-a(x+iy)^2}| = e^{ay^2} e^{-ax^2};$$

donc, en vertu de 15.26*b*, on peut passer, dans le plan des z, de l'axe des x à n'importe quelle droite parallèle pour calculer la transformée de Fourier:

$$\psi(\sigma) = \int_{-\infty}^{\infty} e^{-a(x+iy)^2} e^{-i\sigma(x+iy)}\, dx = \int_{-\infty}^{\infty} e^{-ax^2+ay^2+\sigma y-2aixy-i\sigma x}\, dx =$$

$$= e^{ay^2+\sigma y} \int_{-\infty}^{\infty} e^{-ax^2-ix(2ay+\sigma)}\, dx.$$

Posons $y = -\sigma/(2a)$; alors $ay^2 + \sigma y = -\sigma^2/(4a)$, et l'on a d'après la formule 11.56(2):

$$\psi(\sigma) = e^{-\frac{\sigma^2}{4a}} \int_{-\infty}^{\infty} e^{-ax^2}\, dx = e^{-\frac{\sigma^2}{4a}} \sqrt{\frac{\pi}{a}}.$$

En particulier, pour $\varphi(x) = e^{-x^2/2}$ $(a = 1/2)$, on obtient $\psi(\sigma) = \sqrt{2\pi}\, e^{-\sigma^2/2}$.

15.33. Transformation de Fourier et produit de convolution. Dans 11.48, nous avons appelé *produit de convolution* de deux fonctions $f(x)$ et $g(x)$ définies pour

$-\infty < x < \infty$ la fonction

$$h(x) = f(x) * g(x) = \int_{-\infty}^{\infty} f(\xi)\, g(x-\xi)\, d\xi. \qquad (1)$$

Si $f(x)$ et $g(x)$ sont continues, bornées et absolument intégrables sur $(-\infty, \infty)$, alors $h(x)$ existe pour tout x et est aussi continue, bornée et absolument intégrable sur $(-\infty, \infty)$; de plus, on a

$$\int_{-\infty}^{\infty} h(x)\, dx = \int_{-\infty}^{\infty} f(x)\, dx \int_{-\infty}^{\infty} g(x)\, dx. \qquad (2)$$

Appliquons le théorème sur le produit de convolution en remplaçant $f(x)$ et $g(x)$ par $f(x)\, e^{-i\sigma x}$ et $g(x)\, e^{-i\sigma x}$ respectivement. Le produit de convolution de ces fonctions est

$$\int_{-\infty}^{\infty} f(\xi)\, e^{-i\sigma\xi} g(x-\xi)\, e^{-i\sigma(x-\xi)}\, d\xi = e^{-i\sigma x} \int_{-\infty}^{\infty} f(\xi)\, g(x-\xi)\, d\xi,$$

et l'égalité (2) donne

$$F[f*g] = \int_{-\infty}^{\infty} e^{-i\sigma x} \left\{ \int_{-\infty}^{\infty} f(\xi)\, g(x-\xi)\, d\xi \right\} dx =$$

$$= \int_{-\infty}^{\infty} f(x)\, e^{-i\sigma x}\, dx \cdot \int_{-\infty}^{\infty} g(x)\, e^{-i\sigma x}\, dx = F[f] \cdot F[g].$$

Autrement dit, dans les suppositions imposées plus haut aux fonctions $f(x)$ et $g(x)$, *la transformée de Fourier de leur produit de convolution est le produit de leurs transformées de Fourier.*

15.34. Résolution de l'équation de la chaleur. Trouvons une solution $u(x, t)$ de l'équation de la chaleur $(-\infty < x < \infty,\ t \geqslant 0)$:

$$\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2} \qquad (1)$$

qui coïncide avec la fonction donnée $u_0(x)$ pour $t = 0$. L'interprétation physique du problème posé consiste à déterminer la température du continu homogène unidimensionnel (d'une barre infinie) à tout instant $t > 0$ si l'on connaît sa température à l'instant $t = 0$. Faisons les suppositions suivantes :

1) les fonctions $u(x, t)$, $u_x(x, t)$, $u_{xx}(x, t)$ sont continues et absolument intégrables en x pour $-\infty < x < \infty$ et pour tout $t \geqslant 0$ fixe ;

2) la fonction $u_t(x, t)$ admet, dans tout intervalle $0 \leqslant t \leqslant T$, un majorant intégrable :

$$|u_t(x, t)| \leqslant \Phi(x), \quad \int_{-\infty}^{\infty} \Phi(x)\, dx < \infty.$$

Appliquons à l'équation (1) la transformation de Fourier en la multipliant par $e^{-i\sigma x}$ et en intégrant ensuite en x de $-\infty$ à ∞. En vertu de la condition 2), de 11.45*a* et de 11.47*a*, on peut écrire

$$\int_{-\infty}^{\infty} u_t(x, t)\, e^{-i\sigma x}\, dx = \frac{\partial}{\partial t} \int_{-\infty}^{\infty} u(x, t)\, e^{-i\sigma x}\, dx = v_t(\sigma, t),$$

où

$$v(\sigma, t) = \int_{-\infty}^{\infty} u(x, t)\, e^{-i\sigma x}\, dx$$

est la transformée de Fourier de la solution cherchée $u(x, t)$. D'après la condition 1) et la formule 15.25(1), on a

$$F[u_{xx}(x, t)] = -\sigma^2 F[u] = -\sigma^2 v(\sigma, t).$$

Nous aboutissons à l'équation différentielle ordinaire

$$v_t(\sigma, t) = -\sigma^2 v(\sigma, t)$$

dont on doit trouver une solution qui coïncide, pour $t = 0$, avec

$$v_0(\sigma) = F[u_0(x)] = \int_{-\infty}^{\infty} u_0(x)\, e^{-i\sigma x}\, dx.$$

La solution cherchée a, évidemment, la forme

$$v(\sigma, t) = e^{-\sigma^2 t} v_0(\sigma).$$

Nous savons (cf. 15.32 où l'on doit poser $a = 1/(4t)$) que

$$e^{-\sigma^2 t} = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-\frac{x^2}{4t}}\right].$$

D'après la formule pour la transformée de Fourier d'un produit de convolution (15.33(2)), nous avons

$$v(\sigma, t) = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-\frac{x^2}{4t}}\right] F[u_0] = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-\frac{x^2}{4t}} * u_0(x)\right],$$

et, comme $v(\sigma, t) = F[u(x, t)]$, nous aboutissons à

$$u(x, t) = \frac{1}{2\sqrt{\pi t}}\, e^{-\frac{x^2}{4t}} * u_0(x) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-\frac{\xi^2}{4t}} u_0(x - \xi)\, d\xi.$$

La formule obtenue s'appelle *intégrale de Poisson*. Dans la théorie des équations aux dérivées partielles, on démontre l'unicité de la solution ci-dessus dans une vaste classe de fonctions [11].

§ 15.4. Transformation de Laplace

15.41. Soit $\varphi(x)$ une fonction donnée pour $-\infty < x < \infty$ qui est continue par morceaux et qui devient absolument intégrable après l'avoir multipliée par $e^{-\gamma x}$ avec un γ réel. Alors la transformée de Fourier de la fonction $\varphi(x)$ qui n'existe forcément pas au sens initial de ce terme peut s'avérer existante pour certains s complexes; en particulier,

$$\psi(s) = \int_{-\infty}^{\infty} \varphi(x)\, e^{-isx}\, dx = \int_{-\infty}^{\infty} \varphi(x)\, e^{-ix\sigma} e^{x\tau}\, dx$$

existe sur la droite $\tau = -\gamma$. Nous voyons que, sur cette droite, $\psi(s)$ est la transformée de Fourier de la fonction absolument intégrable $\varphi(x)\, e^{x\tau}$.

Le plus important cas se réalise en l'occurrence dans les conditions

$$\left.\begin{aligned} &|\varphi(x)| < Ce^{ax} && \text{pour } x \geqslant 0, \\ &\varphi(x) = 0 && \text{pour } x < 0. \end{aligned}\right\} \tag{1}$$

Ici la transformée de Fourier

$$\psi(s) = \int_0^{\infty} \varphi(x)\, e^{x\tau} e^{-ix\sigma}\, dx = \int_0^{\infty} \varphi(x)\, e^{-isx}\, dx \tag{2}$$

existe pour tout $\tau < -\alpha$, i.e. dans le demi-plan de la variable complexe $s = \sigma + i\tau$ au-dessous de la droite $\tau = -\alpha$. Effectuons dans la formule (2) le changement de variable $is = p$. Lorsque s parcourt le demi-plan $\operatorname{Im} s < -\alpha$, p parcourt le demi-plan $\operatorname{Re} p > \alpha$. La fonction

$$\Phi(p) = \psi(s) = \int_0^{\infty} \varphi(x)\, e^{-px}\, dx$$

est définie et analytique dans le demi-plan $\operatorname{Re} p > \alpha$; sur toute droite verticale de ce demi-plan elle tend vers zéro lorsque $\operatorname{Im} p \to \pm\infty$, cette convergence étant uniforme sur tout intervalle fermé fini des valeurs $\operatorname{Re} p$. En outre, dans le demi-plan $\operatorname{Re} p > \alpha$ on a l'estimation suivante pour la fonction $\Phi(p)$ $(p = \xi + i\eta)$:

$$|\Phi(p)| \leqslant \int_0^{\infty} |\varphi(x)|\, e^{-\xi x}\, dx \leqslant C \int_0^{\infty} e^{(\alpha-\xi)x}\, dx = \frac{C}{\xi - \alpha}.$$

Il en résulte que la fonction $\Phi(p)$ est bornée dans tout demi-plan $\operatorname{Re} p \geqslant \beta > \alpha$ et qu'elle tend vers zéro lorsque $\xi \to \infty$.

La fonction $\Phi(p)$ s'appelle *transformée de Laplace* de la fonction $\varphi(x)$. Nous voyons que la transformation de Laplace ne diffère de celle de Fourier (considérée dans le domaine complexe) que par une rotation de 90° dans le plan de la variable complexe.

15.42.a. Le théorème simple suivant fournit les conditions suffisantes (mais qui sont loin d'être nécessaires) pour qu'une fonction $\Phi(p)$ donnée soit la transformée de Laplace d'une fonction $\varphi(x)$ vérifiant les conditions 15.41(1).

Théorème. *Soit* $\Phi(p)$ $(p = \xi + i\eta)$ *une fonction satisfaisant aux conditions suivantes*:

1) *elle est analytique dans un demi-plan* $\operatorname{Re} p > \gamma_0 \geqslant 0$;

2) *il existe une constante* C *et une fonction* $B(\eta)$ *positive intégrable sur l'axe* $-\infty < \eta < \infty$ *telles que, pour tous les* $\xi > \gamma_0$, *on ait l'estimation*

$$\left|\Phi(p) - \frac{C}{p}\right| \leqslant B(\eta).$$

Alors $\Phi(p)$ *est la transformée de Laplace d'une fonction* $\varphi(x)$ *continue par morceaux qui est nulle pour* $x < 0$, *continue pour* $x > 0$ *et vérifie l'inégalité*

$$|\varphi(x)| < Ce^{\gamma_0 x}$$

pour $x > 0$.

Démonstration. La fonction C/p est évidemment la transformée de Laplace de la fonction $\varphi_0(x)$ qui vaut 0 pour $x < 0$ et C pour $x > 0$. La fonction $\varphi_0(x)$ satisfait aux exigences du théorème. En la séparant nous pouvons supposer que la fonction $\Phi(p)$ elle-même vérifie, pour $\xi > \gamma_0$, l'inégalité

$$|\Phi(p)| \leqslant B(\eta).$$

Dans ce cas, nous définissons la fonction $\varphi(x)$ par la formule

$$\varphi(x) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \Phi(p)\, e^{px}\, dp \quad (\gamma > \gamma_0). \tag{1}$$

En utilisant la formule de Cauchy (de manière analogue à 15.26*b*) et en se basant sur les propriétés 1), 2) il est aisé de prouver que l'intégrale (1) ne dépend pas de γ. Par ailleurs, on a l'estimation

$$|\varphi(x)| \leqslant \frac{1}{2\pi} \int_{-\infty}^{\infty} |\Phi(\xi + i\eta)|\, e^{\gamma x}\, d\eta \leqslant \frac{B}{2\pi}\, e^{\gamma x}.$$

Pour $x > 0$, en faisant tendre γ vers γ_0 on a

$$|\varphi(x)| \leqslant Ce^{\gamma_0 x};$$

pour $x < 0$, en faisant tendre γ vers $+\infty$, on a $\varphi(x) \equiv 0$.

Si l'on met la formule (1) sous la forme

$$\varphi(x) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \Phi(\xi + i\eta) e^{(\xi+i\eta)x} i\, d\eta = \frac{e^{\xi x}}{2\pi} \int_{-\infty}^{\infty} \Phi(\xi + i\eta) e^{i\eta x}\, d\eta,$$

on voit que $2\pi\varphi(-x)\, e^{\xi x}$ est la transformée de Fourier, par rapport à la variable η, de la fonction absolument intégrable $\Phi(\xi + i\eta)$ (ξ fixe). D'apres la formule d'inversion, on a

$$\Phi(\xi + i\eta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} 2\pi\varphi(-x)\, e^{(\xi+i\eta)x}\, dx = \int_{-\infty}^{\infty} \varphi(x)\, e^{-px}\, dx, \qquad (2)$$

de sorte que $\Phi(\xi + i\eta)$ est effectivement la transformée de Laplace de la fonction $\varphi(x)$.

La formule (1) joue un rôle important dans la théorie de la transformation de Laplace ; elle s'appelle *formule d'inversion de Laplace*.

b. Voyons quelles sont les conditions à imposer à la fonction $\varphi(x)$ pour que sa transformée de Laplace vérifie l'hypothèse du théorème a. Supposons que $\varphi(x)$ soit $m - 1$ fois continûment dérivable et possède la dérivée m-ième continue par morceaux, ces dérivées satisfaisant aux conditions 15.41(1). Alors, en intégrant l'égalité (2) m fois par parties nous obtenons, pour $\xi = \operatorname{Re} p \geqslant \gamma > \alpha$:

$$|\Phi(p)| = \left| \frac{1}{p^m} \int_0^{\infty} \varphi^{(m)}(x)\, e^{-px}\, dx \right| \leqslant \frac{C}{|p|^m} =$$

$$= \frac{C}{|\xi^2 + \eta^2|^{m/2}} \leqslant \frac{Cl}{(\gamma^2 + \eta^2)^{m/2}}. \qquad (3)$$

Nous voyons que l'hypothèse du théorème a est vérifiée si $m = 2$. Donc, l'existence de la dérivée seconde continue par morceaux de la fonction $\varphi(x)$ garantit que l'hypothèse du théorème a soit vérifiée.

15.43. La transformation de Laplace aide souvent à résoudre les équations différentielles, ordinaires ou aux dérivées partielles, qui correspondent à des systèmes non stationnaires ; dans de tels problèmes, la fonction inconnue $f(t)$ est nulle pour $t < 0$ et, pour $t > 0$, doit satisfaire à une équation et à certaines conditions initiales pour $t = 0$.

Considérons d'abord une équation différentielle linéaire à coefficients constants

$$a_0 y^{(n)}(t) + a_1 y^{(n-1)}(t) + \ldots + a_n y(t) = b(t) \qquad (1)$$

complétée par les valeurs données

$$\begin{aligned} y(0) &= y_0, \\ y'(0) &= y_1, \\ &\cdots\cdots\cdots \\ y^{(n-1)}(0) &= y_{n-1}, \end{aligned} \tag{2}$$

$b(t)$ vérifiant les conditions 15.41(1). Multiplions l'équation (1) par e^{-pt} et intégrons en t de 0 à ∞. Désignons par

$$Y(p) = \int_0^\infty y(t)\, e^{-pt}\, dt$$

la transformée de Laplace de la fonction $y(t)$. Alors, en intégrant par parties, nous trouvons

$$\left.\begin{aligned} \int_0^\infty y'(t)\, e^{-pt}\, dt &= y(t)\, e^{-pt}\Big|_0^\infty + p\int_0^\infty y(t)\, e^{-pt}\, dt = \\ &= -y_0 + pY(p), \\ \int_0^\infty y''(t)\, e^{-pt}\, dt &= y'(t)\, e^{-pt}\Big|_0^\infty + p\int_0^\infty y'(t)\, e^{-pt}\, dt = \\ &= -y_1 + p(-y_0 + pY(p)) = \\ &= -y_1 - py_0 + p^2 Y(p), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \int_0^\infty y^{(n)}(t)\, e^{-pt}\, dt &= y^{(n-1)}(t)\, e^{-pt}\Big|_0^\infty + \\ &\quad + p\int_0^\infty y^{(n-1)}(t)\, e^{-pt}\, dt = \\ &= -y_{n-1} + p(-y_{n-2} - py_{n-3} - \dots \\ &\quad \dots - p^{n-2}y_0 + p^{n-1}Y(p)) = \\ &= -y_{n-1} - py_{n-2} - \dots \\ &\quad \dots - p^{n-1}y_0 + p^n Y(p). \end{aligned}\right\} \tag{3}$$

En multipliant chacune des équations (3) par le coefficient a_k correspondant et en additionnant nous avons l'équation de la forme

$$R_0(p) + R(p)\, Y(p) = B(p),$$

où $R_0(p)$ est un polynôme en p de degré au plus égal à $n-1$, $R(p)$ un polynôme en p de degré n, $B(p)$ la transformée de Laplace de la

fonction $b(t)$. Pour la fonction inconnue $Y(p)$, on obtient ainsi une équation purement algébrique. En la résolvant on trouve

$$Y(p) = \frac{B(p) - R_0(p)}{R(p)}.$$

La fonction $\frac{R_0(p)}{R(p)}$ vérifie l'hypothèse du théorème 15.42a; notamment, si l'on pose

$$C = \lim_{p \to \infty} \frac{pR_0(p)}{R(p)},$$

alors

$$\frac{R_0(p)}{R(p)} - \frac{C}{p} = \frac{pR_0(p) - CR(p)}{pR(p)}$$

est une fraction rationnelle en p, le degré $n+1$ du dénominateur dépassant d'au moins deux unités celui du numérateur car, en vertu de la définition de C, les termes de degré n du numérateur se réduisent.

Quant à la fonction $\frac{B(p)}{R(p)}$, elle vérifie ou non l'hypothèse du théorème 15.42a; cela dépend de la nature de la fonction $B(p)$. Si le degré du polynôme $R(p)$ est supérieur à 1, il suffit pour cela, $B(p)$ étant bornée, que la fonction $b(t)$ satisfasse aux conditions 15.41(1); si le degré de $R(p)$ est 1, alors l'inégalité 15.42(3) montre qu'il suffit que la fonction $b(t)$ possède la dérivée continue par morceaux satisfaisant avec $b(t)$ aux conditions 15.41(1).

Si la fonction $\frac{B(p)}{R(p)}$ vérifie également l'hypothèse du théorème 15.42a, alors en l'appliquant nous avons pour la solution $y(t)$ la formule

$$y(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{B(p) - R_0(p)}{R(p)} e^{pt}\, dp. \tag{4}$$

Si la fonction $B(p)$ est analytiquement prolongeable dans tout le plan des p (avec des singularités isolées), alors on calcule d'ordinaire l'intégrale (4) à l'aide de l'intégration le long d'un contour et de la théorie des résidus, comme nous l'avons fait en calculant l'intégrale de Fourier de fonctions rationnelles. Remarquons que, pour $t > 0$, la fonction e^{pt} est bornée dans le demi-plan à gauche ($\operatorname{Re} p < \gamma$) et ne l'est pas dans l'autre demi-plan; il faut donc construire les demi-circonférences faisant partie du contour à gauche de la droite $\operatorname{Re} p = \gamma$, et non à droite. En tant que γ on peut choisir n'importe quel nombre, pourvu que toutes les singularités de la fonction $R(p)$ soient à gauche de la droite $\operatorname{Re} p = \gamma$.

15.44. Exemple. Considérons une équation du deuxième ordre

$$a_0 y'' + a_1 y' + a_2 y = b \sin kt, \quad y_0 = 0, \quad y_1 = 0$$

dont les racines caractéristiques (13.17) sont les nombres complexes conjugués (uon réels) $\lambda = \alpha + i\beta$, $\bar{\lambda} = \alpha - i\beta$, où $\alpha < 0$.

En électricité, une telle équation décrit les oscillations forcées dans un circuit contenant une résistance, une self et une capacité et soumis à une f.é.m. de fréquence k. Le passage à la transformée de Laplace amène à l'équation

$$(a_0p^2 + a_1p + a_2)\,Y(p) = \int_0^\infty b \sin kte^{-pt}dt = \frac{bk}{k^2 + p^2}.$$

En la résolvant on trouve

$$Y(p) = \frac{bk}{(a_0p^2 + a_1p + a_2)(k^2 + p^2)}.$$

D'après la formule d'inversion, on a

$$y(t) = \frac{bk}{2\pi i}\int_{\gamma - i\infty}^{\gamma + i\infty} \frac{e^{pt}\,dp}{(a_0p^2 + a_1p + a_2)(k^2 + p^2)}.$$

Posons

$$f(p) = \frac{e^{pt}}{(a_0p^2 + a_1p + a_2)(k^2 + p^2)}.$$

Le dénominateur possède quatre racines simples aux points $\pm ik$ et $\alpha \pm i\beta$. En tant que γ on peut choisir n'importe quel nombre positif. Pour calculer l'intégrale, complétons la droite $\operatorname{Re} p = \gamma$ par une demi-circonférence de rayon suffisamment grand dans le demi-plan à gauche (fig. 15.2); alors on a, d'après le théorème des résidus 10.43:

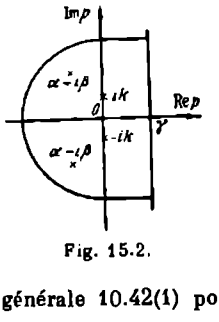


Fig. 15.2.

$$y(t) = bk\{\operatorname{Res} f(p)|_{p=ik} + \operatorname{Res} f(p)|_{p=-ik} + \operatorname{Res} f(p)|_{p=\alpha+i\beta} + \operatorname{Res} f(p)|_{p=\alpha-i\beta}\}.$$

Chaque résidu est calculé selon la formule générale 10.42(1) pour les pôles simples:

$$\operatorname{Res}\frac{A(p)}{B(p)}\Big|_{p=p_0} = \frac{A(p_0)}{B'(p_0)}.$$

Définitivement, on a

$$y(t) = bk\left[\frac{e^{(\alpha+i\beta)t}}{(\lambda^2 + k^2)\,2i\beta a_0} - \frac{e^{(\alpha-i\beta)t}}{(\bar{\lambda}^2 + k^2)\,2i\beta a_0} + \frac{e^{ikt}}{(-a_0k^2 + a_1ik + a_2)\,2ik} - \frac{e^{-ikt}}{(-a_0k^2 - a_1ik + a_2)\,2ik}\right].$$

Le processus résultant est la superposition d'une oscillation périodique dont la fréquence est celle de la force extérieure et d'une oscilla-

tion amortie dont la fréquence est la fréquence propre du système; la vitesse d'amortissement est déterminée par la quantité α, i.e. par l'abscisse des racines caractéristiques.

Pour $\alpha = 0$, $\beta = k$, on a une résonance. L'équation initiale prend la forme

$$y'' + k^2 y = b \sin kt$$

avec pour solution

$$y(t) = \frac{bk}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{e^{pt}\,dp}{(p^2 - k^2)^2}.$$

Les points $p = \pm ik$ sont les pôles de multiplicité 2 de la fonction sous le signe somme. En calculant les résidus d'après la formule 10.42(2) on trouve

$$y(t) = bk\left[e^{ikt}\left(-\frac{t^2}{4k^2} + \frac{1}{4ik^3}\right) + e^{-ikt}\left(-\frac{1}{4k^2} - \frac{1}{4ik^3}\right)\right] =$$

$$= \frac{bt}{2k}\cos kt - \frac{b}{2k^2}\sin kt,$$

ce qui représente une oscillation d'amplitude indéfiniment croissante.

15.45. Les mêmes méthodes s'appliquent aux équations aux dérivées partielles.

Si l'on applique la transformation de Laplace, une équation différentielle ordinaire se transforme en une équation algébrique par rapport à la fonction inconnue, tandis qu'une équation contenant les dérivées non seulement en t, mais aussi en x, y, . . ., ne renferme plus de dérivées en t, les dérivées en x, y, . . . étant conservées.

En tant qu'exemple, considérons l'équation de la chaleur $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ dans un intervalle fini $0 \leqslant x \leqslant l$ avec les conditions $u_x(0, t) = 0$, $u(l, t) = u_1$, $u(x, 0) = u_0$. Du point de vue physique, cela veut dire que la chaleur ne s'écoule pas par l'extrémité $x = 0$, qu'à l'extrémité $x = l$ on maintient la température constante u_1 par un apport de chaleur de l'extérieur (pour $t > 0$) et qu'à l'instant initial la température est constante et vaut u_0.

Appliquons la transformation de Laplace en t, i.e. passons de la fonction $u(x, t)$ à la fonction

$$v(x, p) = \int_0^\infty e^{-pt} u(x, t)\,dt.$$

Pour $v(x, p)$, nous obtenons l'équation

$$\frac{d^2 v(x, p)}{dx^2} - pv(x, p) = -u_0$$

avec les conditions

$$v_x(0, p) = 0, \quad v(l, p) = \frac{u_1}{p}.$$

Cette équation du 2-ième ordre a pour solution

$$v(x,\ p)=\frac{u_0}{p}+\frac{u_1-u_0}{p}\,\frac{\operatorname{ch} x\sqrt{p}}{\operatorname{ch} l\sqrt{p}},$$

d'où

$$u(x,\ t)=u_0+\frac{u_1-u_0}{2\pi i}\int\limits_{\gamma-i\infty}^{\gamma+i\infty}e^{pt}\,\frac{\operatorname{ch} x\sqrt{p}}{\operatorname{ch} l\sqrt{p}}\,\frac{dp}{p}. \qquad (1)$$

La fonction sous le signe somme est une fonction univoque de p avec les pôles $p_0=0$ et $p_n=-\frac{\pi^2}{l^2}\left(n-\frac{1}{2}\right)^2$ $(n=1,\ 2,\ \ldots)$. Nous montrerons que l'intégrale

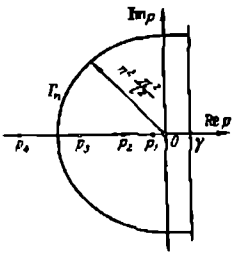


Fig. 15.3.

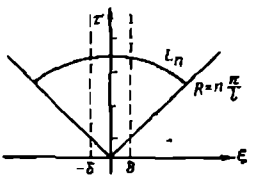


Fig. 15.4.

est égale à la somme (infinie) des résidus de la fonction correspondante en tous ces pôles. Pour le faire, considérons dans le demi-plan à gauche la demi-circonférence Γ_n centrée à l'origine des coordonnées et de rayon $n^2\pi^2/l^2$ (fig. 15.3); elle passe entre deux pôles voisins, et nous montrerons que le rapport $\frac{\operatorname{ch} x\sqrt{p}}{\operatorname{ch} l\sqrt{p}}$ est borné sur toute la demi-circonférence; alors, en vertu du lemme de Jordan 11.32*d*, l'intégrale sur Γ_n tendra vers zéro pour $n\to\infty$, et toute l'intégrale (1) se réduira, comme d'ordinaire, à la somme des résidus.

Au lieu de considérer le rapport $\frac{\operatorname{ch} x\sqrt{p}}{\operatorname{ch} l\sqrt{p}}$ sur la demi-circonférence Γ_n où $|p|=n^2\pi^2/l^2$, on peut remplacer $\sqrt{p}$ par ζ, p par ζ^2 et considérer le rapport $\frac{\operatorname{ch} x\zeta}{\operatorname{ch} l\zeta}$ sur le quart de la circonférence L_n de rayon $n\pi/l$ et avec l'angle polaire qui varie de $\pi/4$ à $3\pi/4$ (fig. 15.4). Posons $\zeta=\xi+i\tau$, nous avons $\tau>0$, $|\xi|<\tau$ et

$$\left|\frac{\operatorname{ch} x\zeta}{\operatorname{ch} l\zeta}\right|^2=\left|\frac{\operatorname{ch} x(\xi+i\tau)}{\operatorname{ch} l(\xi+i\tau)}\right|^2=\left|\frac{\operatorname{ch} x\xi\cos x\tau+i\operatorname{sh} x\xi\sin x\tau}{\operatorname{ch} l\xi\cos l\tau+i\operatorname{sh} l\xi\sin l\tau}\right|^2=$$

$$=\frac{\operatorname{ch}^2 x\xi\cos^2 x\tau+\operatorname{sh}^2 x\xi\sin^2 x\tau}{\operatorname{ch}^2 l\xi\cos^2 l\tau+\operatorname{sh}^2 l\xi\sin^2 l\tau}<\frac{\operatorname{ch}^2 l\xi}{\operatorname{ch}^2 l\xi\cos^2 l\tau+\operatorname{sh}^2 l\xi\sin^2 l\tau}. \qquad (2)$$

Si $|\xi|<\delta$, alors nous avons $|\tau-n\pi/l|<\varepsilon$ sur la circonférence L_n pour n suffisamment grand, donc $\cos^2 l\tau<1-\eta$, où ε, η sont arbitrairement petits;

par conséquent,

$$\left|\frac{\operatorname{ch} x\zeta}{\operatorname{ch} l\zeta}\right|^2 \leqslant \frac{\operatorname{ch}^2 l\xi}{(1-\eta)\operatorname{ch}^2 l\xi} = \frac{1}{1-\eta}. \tag{3}$$

Si $|\xi| \geqslant \delta$, alors nous remplaçons dans le dénominateur du dernier membre de (2) $\operatorname{ch}^2 l\xi$ par $\operatorname{sh}^2 l\xi$ en obtenant

$$\left|\frac{\operatorname{ch} x\zeta}{\operatorname{ch} l\zeta}\right|^2 \leqslant \frac{\operatorname{ch}^2 l\xi}{\operatorname{sh}^2 l\xi} = \operatorname{coth}^2 l\xi \leqslant \operatorname{coth}^2 l\delta. \tag{4}$$

Il résulte de (3) et (4) que le rapport $\left|\frac{\operatorname{ch} x\sqrt{p}}{\operatorname{ch} l\sqrt{p}}\right|$ est borné sur les circonférences mentionnées par une constante fixe. Donc, comme nous l'avons déjà dit, l'intégrale se réduit à la somme des résidus. Le résidu au pôle $p=0$ vaut 1. Le résidu au pôle $p_n = -\frac{\pi^2}{l^2}\left(n-\frac{1}{2}\right)^2$ est égal, on le calculera aisément, à

$$\frac{(-1)^n \cdot 4}{\pi(2n-1)} e^{-\frac{\pi^2}{l^2}\left(n-\frac{1}{2}\right)^2 t} \cos\left(n-\frac{1}{2}\right)\frac{\pi x}{l}.$$

En fin de compte, nous obtenons la solution sous la forme de la somme d'une série :

$$u(x, t) = u_0 + \frac{4}{\pi}(u_1 - u_0) \sum_{n=1}^{\infty} \frac{(-1)^n}{2n-1} e^{-\frac{\pi^2}{l^2}\left(n-\frac{1}{2}\right)^2 t} \cos\left(n-\frac{1}{2}\right)\frac{\pi x}{l}.$$

§ 15.5 *. Classes quasi analytiques de fonctions

15.51. La transformation de Laplace s'applique avec succès à des problèmes d'ordre théorique. Le théorème fondamental de la théorie des classes quasi analytiques de fonctions en est un exemple *).

On sait qu'une fonction $f(x)$ d'une variable réelle x, si elle est indéfiniment dérivable au voisinage d'un point x_0, n'est pas nécessairement analytique, i.e. ne se développe pas forcément en une série de Taylor au voisinage de ce point. Mais si les dérivées successives de la fonction $f(x)$ ne croissent pas trop rapidement, notamment satisfont aux conditions

$$\max_{|x-x_0|<\delta} |f^{(n)}(x)| \leqslant CM^n n!, \tag{1}$$

alors, comme nous l'avons vu dans 8.52, la fonction $f(x)$ est analytique au voisinage du point x_0.

En appliquant la formule de Cauchy 10.34(1) pour les dérivées d'une fonction analytique on peut prouver facilement que, inversement, l'analyticité d'une fonction $f(x)$ au voisinage d'un point x_0 entraîne les inégalités (1). Soit $m_0, m_1, \ldots, m_n, \ldots$ une suite quelconque de nombres positifs. Introduisons la classe $C_{\{m_n\}}$ des fonctions $f(x)$ définies sur l'axe $-\infty < x < \infty$ et vérifiant les

*) D'après [6].

inégalités

$$|f^{(n)}(x)| \leqslant CM^n m_n \quad (n=0,\ 1,\ 2,\ \ldots),$$

où les constantes C et M peuvent dépendre du choix de la fonction $f(x)$. Si les nombres m_n croissent plus rapidement que $n!$, alors la classe $C_{\langle m_n \rangle}$ peut contenir aussi bien des fonctions non analytiques. Cependant, Denjoy a montré en 1921 qu'il existe des classes $C_{\langle m_n \rangle}$ contenant des fonctions non analytiques mais possédant tout de même la propriété d'unicité : *si deux fonctions $f(x)$ et $g(x)$ contenues dans la classe $C_{\langle m_n \rangle}$ coïncident en un point x_0, de même que toutes leurs dérivées respectives, alors elles sont identiques.* Pour les fonctions analytiques, cette propriété est bien connue (elle résulte de 10.39*e*).

15.52. Les classes $C_{\langle m_n \rangle}$ dans lesquelles la coïncidence de deux fonctions et de leurs dérivées respectives de tous les ordres en un point implique l'identité de ces fonctions sont appelées *classes quasi analytiques*. En 1926, Carleman donne la description complète des classes quasi analytiques ; en 1930, Ostrovski propose un énoncé plus simple. Pour comprendre l'énoncé du théorème de Carleman-Ostrovski, il nous faut accomplir quelques constructions préalables. Nous supposons que la suite m_n croît pour $n \to \infty$ plus rapidement que toute fonction de la forme r^n, où $r > 0$ (un peu plus bas nous montrerons que dans le cas contraire le problème est très simple). Alors, pour tout $r > 0$, la suite r^n/m_n tend vers 0 lorsque $n \to \infty$, donc est bornée. Dans ce qui suit, le rôle principal échoit à la fonction

$$T(r) = \sup_{n \geqslant 0} \frac{r^n}{m_n}.$$

La fonction $T(r)$ admet une interprétation géométrique utile. Considérons, dans le demi-plan à droite du plan des x, y, la suite des points de coordonnées $x_n = n$, $y_n = -\ln m_n$ (*points de Valiron*). Comme $r^n/m_n \to 0$ pour $n \to \infty$, on a $n \ln r - \ln m_n \to -\infty$ de sorte que, quels que soient r et $b > 0$, il n'y a qu'un nombre fini de points de Valiron vérifiant l'inégalité

$$n \ln r - \ln m_n \geqslant b. \tag{1}$$

A l'équation $x \ln r + y = b$ ou

$$y = -x \ln r + b$$

il correspond, dans le demi-plan à droite des x, y, une demi-droite de coefficient angulaire $-\ln r$ qui coupe l'axe des y au point b. L'inégalité (1) montre qu'au-dessus d'une telle demi-droite il n'y a qu'un nombre fini de points de Valiron. Par conséquent, pour tout r on peut trouver un $b = b(r)$ tel qu'aucun point de Valiron ne se trouve au-dessus de la droite

$$y = -x \ln r + b(r),$$

tandis que sur la droite même il en existe au moins un (fig. 15.5). Une telle demi-droite sera appelée *demi-droite de Valiron*.

Pour tous $b = b(r)$ et n, nous avons par construction

$$-n \ln r + b(r) \geqslant -\ln m_n,$$

de sorte que

$$b(r) \geqslant \sup_n \{n \ln r - \ln m_n\} = \sup_n \ln \frac{r^n}{m_n}. \tag{2}$$

Or, étant donné que l'inégalité (2) devient l'égalité pour au moins un numéro n, nous avons en fait

$$b(r) = \sup_n \ln \frac{r^n}{m_n},$$

donc

$$b(r) = \ln T(r). \tag{3}$$

L'interprétation géométrique de la fonction $\ln T(r)$ nous aidera à établir certaines de ses propriétés. Tout d'abord, il résulte de la définition de la fonction $b(r)$ que $b(r)$ est une fonction croissante de r: si, pour $r_2 > r_1$, on avait $b(r_2) < b(r_1)$, alors toute la demi-droite $y = -x \ln r_2 + b(r_2)$ passerait au-dessous de la demi-droite $y = -x \ln r_1 + b(r_1)$, ce qui contredirait le fait que celle-ci porte au moins un point de Valiron. Ensuite, on peut toujours construire une demi-droite de Valiron d'après la valeur donnée de b ($> \inf b(r)$), en considérant la famille de toutes les demi-droites coupant l'axe des ordonnées au point b et en raisonnant comme ci-dessus. Cela veut dire que la fonction croissante $b(r) = \ln T(r)$ prend *toutes* les valeurs ($> \inf b(r)$), donc est continue (5.36). (On pourrait également prouver qu'elle est une fonction linéaire par morceaux de $\ln r$, mais on n'en a pas besoin.)

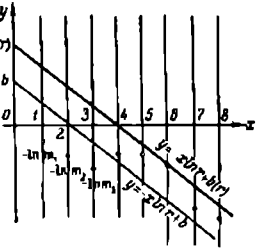


Fig. 15.5.

Nous sommes maintenant en mesure de donner l'énoncé d'Ostrovski du théorème de Carleman:

Théorème. *Posons*

$$T(r) = \sup_{n \geqslant 0} \frac{r^n}{m_n}. \tag{4}$$

Alors, pour que la classe $C_{\{m_n\}}$ *soit quasi analytique, il faut et il suffit que*

$$\int_1^\infty \frac{\ln T(r)}{r^2}\, dr = \infty. \tag{5}$$

Soit, par exemple, $m_n = (n!)^\alpha$, où α est fixe. Alors il est aisé de voir, en se servant de la formule de Stirling 11.57*b*, que

$$T(r) \sim r^{1-\alpha}$$

et que l'intégrale (5) converge donc pour $\alpha > 1$ et diverge pour $\alpha \leqslant 1$. Il résulte alors du théorème de Carleman que la classe $C_{\{(n!)^\alpha\}}$ est quasi analytique pour $\alpha \leqslant 1$ (comme nous l'avons vu plus haut, elle est même formée de fonctions analytiques) et ne l'est pas pour $\alpha > 1$.

Il existe des classes quasi analytiques comprenant entre autres des fonctions non seulement analytiques. On peut prouver que la fonction $f(x) = \sum T^{-1}(n) \cos nx$ appartient à la classe $C_{\{m_n\}}$ et n'est pas analytique si $\sqrt[n]{m_n}/n \to \infty$; donc, pour $m_n = n! \ln^n n$ par exemple, la classe quasi analytique $C_{\{m_n\}}$ renferme des fonctions non analytiques.

15.53. Dans ce qui suit (15.53-15.56) nous démontrons le théorème de Carleman-Ostrovski énoncé dans 15.52.

Dans le présent numéro, en utilisant la transformation de Laplace nous réduisons le problème de caractérisation des classes quasi analytiques à un problème sur les fonctions analytiques dans un demi-plan.

Supposons que la classe $C_{\{m_n\}}$ ne soit pas quasi analytique. Cela signifie qu'il y existe deux fonctions $f(x)$ et $g(x)$ qui coïncident, de même que leurs dérivées respectives, en $x = x_0$, sans coïncider partout. Sans restreindre la généralité, on peut poser $x_0 = 0$ et $f(x) \neq g(x)$ pour $x > 0$; on peut toujours satisfaire à ces conditions en faisant une translation et en remplaçant x par $-x$, c'est-à-dire en effectuant les opérations réalisables dans la classe $C_{\{m_n\}}$. Considérons ensuite la fonction $\varphi(x)$ qui est nulle pour $x < 0$ et $f(x) - g(x)$ pour $x \geqslant 0$; il est évident qu'elle appartient également à la classe $C_{\{m_n\}}$. Comme cette fonction est nulle pour $x < 0$ et bornée pour $x > 0$, elle possède une transformée de Laplace

$$\Phi(p) = \int_0^\infty \varphi(x) e^{-px} dx \tag{1}$$

qui est analytique dans le demi-plan $\operatorname{Re} p > 0$.

Dégageons quelques autres propriétés de la fonction $\Phi(p)$. En intégrant n fois par parties dans (1) nous obtenons

$$p^n \Phi(p) = \int_0^\infty \varphi^{(n)}(x) e^{-px} dx,$$

d'où l'estimation

$$|p^n \Phi(p)| \leqslant C M^n m_n \int_0^\infty e^{-px} dx = C M^n m_n \frac{1}{|p|} \leqslant C_\gamma M^n m_n$$

pour $|p| > \gamma > 0$. Inversement, soit $\Phi(p) \not\equiv 0$ une fonction analytique donnée dans un demi-plan quelconque $\operatorname{Re} p > \gamma_0 > 0$ et satisfaisant aux inégalités de la forme

$$|p^n \Phi(p)| \leqslant CM^n m_n \quad (n = 0, 1, 2, \ldots).$$

Il est évident que $\Phi(p)/p^2$ vérifie l'hypothèse du théorème 15.42*a*; en tant que majorant intégrable exigé par la condition 2), on peut choisir, par exemple, $Cm_0 \frac{1}{|\gamma_0 + t\eta|^2}$. En vertu du théorème cité, la fonction définie par l'égalité $(\gamma > \gamma_0)$

$$\varphi(x) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{\Phi(p)}{p^2} e^{px}\, dp \tag{2}$$

est nulle pour $x < 0$. Puisque $\Phi(p) \not\equiv 0$, nous avons encore $\varphi(x) \not\equiv 0$ pour $x > 0$. De plus, $\varphi(x)$ a les dérivées de tous les ordres et on a

$$|(\varphi(x)\, e^{-\gamma_0 x})^{(n)}| = \frac{1}{2\pi} \left| \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{\Phi(p)}{p^2} (p - \gamma_0)^n e^{(p - \gamma_0)x}\, dp \right| \leqslant$$

$$\leqslant \frac{CM^n m_n}{2\pi} \int_{\gamma - i\infty}^{\gamma + i\infty} \left| \frac{p - \gamma_0}{p} \right|^n \frac{|dp|}{|p^2|} \leqslant \frac{C}{2\pi} M^n m_n \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{|dp|}{|p|^2} = C' M^n m_n.$$

Nous voyons que $\varphi(x)\, e^{-\gamma_0 x}$ appartient à la classe $C_{\langle m_n \rangle}$. Comme $\varphi(x) = 0$ pour $x < 0$ et $\varphi(x) \not\equiv 0$, la classe $C_{\langle m_n \rangle}$ n'est pas quasi analytique. Donc, *le problème de la quasi-analyticité d'une classe* $C_{\langle m_n \rangle}$ *donnée est équivalent au problème de l'existence d'une fonction* $\Phi(p) \not\equiv 0$ *analytique dans le demi-plan* $\operatorname{Re} p > \gamma_0$ *et vérifiant les inégalités*

$$|p^n \Phi(p)| \leqslant CM^n m_n \quad (n = 0, 1, 2, \ldots)$$

(« problème de Watson »).

15.54. Par l'inversion $p = 2\gamma/s$, le demi-plan $\operatorname{Re} p > \gamma$ est transformé en le cercle $|s - 1| < 1$, et le problème de Watson est réduit au problème suivant : quelles sont les conditions à imposer à une suite m_n pour qu'il existe dans le cercle $|s - 1| < 1$ une fonction analytique $F(s) \not\equiv 0$ vérifiant les inégalités

$$|F(s)| \leqslant CM^n m_n |s|^n ? \tag{1}$$

Par exemple, il n'existe pas de telle fonction si $m_n \leqslant C_1 r_0^n$ pour une suite $n_1, n_2, \ldots$ de numéros. En effet, si

$$|F(s)| \leqslant CM^n C_1 r_0^n |s|^n = CC_1 (Mr_0 |s|)^n \quad (n = n_1, n_2, \ldots),$$

alors en choisissant $|s| < 1/(Mr_0)$ et en passant à la limite pour $n_k \to \infty$, on a $F(s) \equiv 0$ contrairement à la supposition. Ainsi, avec

les suppositions faites sur la suite m_n, la classe $C_{\{m_n\}}$ est quasi analytique. Par ailleurs, dans le présent cas nous avons

$$T(r) = \sup_n \frac{r^n}{m_n} = \infty \text{ pour } r > r_0$$

et la condition (1) est évidemment remplie. Comme nous l'avons dit dans 15.52, en cas de $m_{nk} \leqslant C_k r_0^{nk}$ $(k = 1, 2, \ldots)$ le problème est bien simple.

Revenons au cas général et supposons qu'il existe une fonction $F(s)$ vérifiant les conditions demandées. On peut trouver un ρ tel que $F(\rho) \neq 0$, $|F(\rho + \rho e^{i\theta})| < 1$ pour tous les θ réels et que la fonction $F(s)$ a sur la circonférence $s = \rho + \rho e^{i\theta}$ un seul zéro pour $s = 0$. Toutes les constructions ultérieures sont réalisées dans le cercle $|s - \rho| \leqslant \rho$. En vertu des inégalités (1), on a

$$|F(\rho + \rho e^{i\theta})| \leqslant CM^n m_n \rho^n |1 + e^{i\theta}|^n = CM^n m_n \left|2\rho \cos\frac{\theta}{2}\right|^n.$$

En minimisant en n le deuxième membre, on obtient

$$|F(\rho + \rho e^{i\theta})| \leqslant \frac{C}{\max\limits_n \dfrac{1}{M^n m_n \left|2\rho \cos\dfrac{\theta}{2}\right|^n}}$$

et, d'après la définition de la fonction $T(r)$:

$$|F(\rho + \rho e^{i\theta})| \leqslant \frac{C}{T\left(\dfrac{1}{2M\rho\left|\cos\dfrac{\theta}{2}\right|}\right)},$$

d'où

$$\ln|F(\rho + \rho e^{i\theta})| \leqslant \ln C - \ln T\left(\frac{1}{2M\rho\left|\cos\frac{\theta}{2}\right|}\right).$$

Dans la théorie des fonctions analytiques il y a le théorème suivant (sa démonstration sera exposée dans 15.56) : *si une fonction $\Phi(z)$ est analytique dans le cercle $|z - z_0| < h$, n'est pas nulle pour $z = z_0$, ne dépasse pas en module l'unité, est continue dans le cercle fermé $|z - z_0| \leqslant h$ et possède un seul zéro sur la circonférence $|z - z_0| = h$, alors l'intégrale*

$$-\int_0^{2\pi} \ln|\Phi(z_0 + he^{i\theta})|\, d\theta$$

a une valeur finie.

En appliquant ce théorème à la fonction $\Phi(z) = F(z)$ on voit que la fonction

$$\ln T\left(\frac{1}{2M\rho\left|\cos\frac{\theta}{2}\right|}\right) \leqslant \ln C - \ln|F(\rho + \rho e^{i\theta})|$$

possède elle aussi une intégrale finie par rapport à θ de zéro à 2π. Si l'on fait la substitution

$$2M\rho\left|\cos\frac{\theta}{2}\right| = \frac{1}{r},$$

on aboutit à la convergence de l'intégrale

$$\int_a^\infty \frac{\ln T(r)}{r^2} \frac{1}{\sqrt{M^2\rho^2 - \frac{1}{4r^2}}}\, dr,$$

donc aussi de l'intégrale

$$\int_a^\infty \frac{\ln T(r)}{r^2}\, dr. \tag{2}$$

Définitivement, *si la classe $C_{\{m_n\}}$ n'est pas quasi analytique, alors l'intégrale* (2) *converge.* Ceci démontre la suffisance de la condition de Carleman dans 15.52.

15.55. En abordant la démonstration de la nécessité de la condition de Carleman supposons que l'intégrale 15.54(2) converge. Il en est alors de même de l'intégrale

$$\int_0^{2\pi} \ln T\left(\frac{1}{2M\rho\left|\cos\frac{\theta}{2}\right|}\right) d\theta,$$

donc on peut construire l'intégrale de Poisson

$$G(re^{i\varphi}) = \frac{1}{2\pi}\int_0^{2\pi} \ln T\left(\frac{1}{2\left|\cos\frac{\theta}{2}\right|}\right) \frac{1-r^2}{1-2r\cos(\theta-\varphi)+r^2}\, d\theta$$

qui est une fonction harmonique dans le cercle $r < 1$. Posons $G(s-1) = P(s)$ et désignons par $Q(s)$ la fonction harmonique conjuguée (14.49*c*) dans le cercle $|s-1| < 1$. Soit, ensuite,

$$F(s) = e^{-[P(s)+iQ(s)]}.$$

Alors la fonction $F(s)$ vérifie les inégalités

$$|F(s)| \leqslant m_n |s|^n \quad (n = 0, 1, 2, \ldots) \tag{1}$$

En effet, les inégalités (1) sont équivalentes aux inégalités

$$e^{-P(s)} \leqslant m_n |s|^n \quad (n = 0,\ 1,\ 2,\ \ldots)$$

ou bien à

$$-G(s) = -P(s+1) \leqslant \ln m_n + n \ln (s+1). \tag{2}$$

Les deux termes dans le dernier membre peuvent être représentés comme intégrales de Poisson

$$\ln m_n = \frac{1}{2\pi} \int_0^{2\pi} \frac{\ln m_n (1-r^2)}{1-2r\cos(\theta-\varphi)+r^2}\, d\theta,$$

$$n \ln |s+1| = \frac{1}{2\pi} \int_0^{2\pi} \frac{n \ln |e^{i\theta}+1|\,(1-r^2)}{1-2r\cos(\theta-\varphi)+r^2}\, d\theta,$$

après quoi l'inégalité (2) qui est à démontrer s'écrit comme suit

$$\int_0^{2\pi} \ln \left\{ T\left(\frac{1}{2\left|\cos\frac{\theta}{2}\right|}\right) m_n |1+e^{i\theta}|^n \right\} \frac{1-r^2}{1-2r\cos(\theta-\varphi)+r^2}\, d\theta \geqslant 0. \tag{3}$$

Or $|1+e^{i\theta}| = 2\left|\cos\frac{\theta}{2}\right|$; comme

$$T(r) = \sup_{n \geqslant 0} \frac{r^n}{m_n}$$

on a, pour tout n pris à part,

$$T(r) \geqslant \frac{r^n}{m_n}, \quad T(r)\, m_n r^{-n} \geqslant 1,$$

donc la fonction sous le signe somme dans (3) est non négative. Il en résulte que l'inégalité (3) est juste; par conséquent, l'inégalité (1) l'est aussi, et la classe $C_{\langle m_n \rangle}$ s'avère non quasi analytique d'après 15.53. Cela achève la démonstration du théorème de Carleman.

15.56. Ici nous démontrons le théorème utilisé dans 15.54.

T h é o r è m e. *Si une fonction $f(z)$ est analytique dans un cercle $|z - z_0| < h$, n'est pas nulle pour $z = z_0$, ne dépasse pas en module le nombre 1, est continue dans le cercle fermé $|z - z_0| \leqslant h$ et possède un seul zéro z^* sur la circonférence $|z - z_0| = h$, alors l'intégrale*

$$-\int_0^{2\pi} \ln |f(z_0 + he^{i\theta})|\, d\theta$$

est finie.

Démonstration. Sans restreindre la généralité on peut poser $z_0 = 0$, $h = 1$, $z^* = 1$. Dans le cercle $|z| \leqslant r < 1$ la fonction $f(z)$ est analytique et ne peut y avoir qu'un nombre fini de zéros $z_1, \ldots, z_m$; on paut supposer qu'il n'y ait pas de zéros sur la circonférence $|z| = r$. Considérons le contour fermé C montré fig. 15.6 qui est composé par les arcs de la circonférence $|z| = r$ parcourue dans le sens positif, par les circonférences C_k $(k = 1, 2, \ldots, m)$ de rayon très petit ε parcourues dans le sens négatif et par les courbes $L_k = [z_k'', z_k']$ liant les arcs mentionnés et parcourues chacune deux fois dans les sens opposés. La fonction $\ln f(z)$ est analytique à l'intérieur du contour C, et la valeur $\ln f(0)$ peut être représentée par la formule de Cauchy:

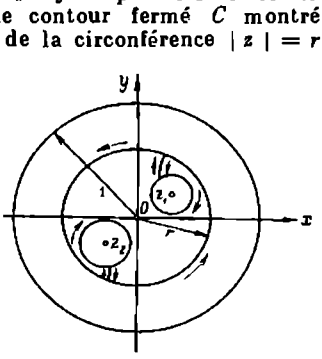


Fig. 15.6.

$$\ln f(0) = \frac{1}{2\pi i} \int_C \ln f(z) \frac{dz}{z}. \qquad (1)$$

Considérons la partie du contour C formée de la circonférence C_j de rayon ε et de centre au point z_j qui est parcourue dans le sens négatif. La partie de l'intégrale (1) prise le long de la circonférence C_j a la forme

$$\frac{1}{2\pi i} \int_{C_j} \ln f(z) \frac{\varepsilon e^{i\theta}\, d\theta}{z_j + \varepsilon e^{i\theta}}. \qquad (2)$$

Si k_j est la multiplicité de la racine z_j, alors $f(z) = (z - z_j)^{k_j} f_j(z)$, où $f_j(z_j) \neq 0$, et l'on a

$$|\ln f(z)| = |\ln (z - z_j)^{k_j} f_j(z)| =$$
$$= |k_j \ln (z - z_j) + \ln f_j(z)| \leqslant k_j |\ln |z - z_j| + 2\pi| +$$
$$+ |\ln f_j(z)| \leqslant k_j |\ln \varepsilon| + C_1.$$

En raison de cette estimation la fonction sous le signe somme dans (2) devient aussi petite que l'on veut pour $\varepsilon \to 0$; par conséquent, toutes les intégrales le long des circonférences C_j tendent vers zéro lorsque $\varepsilon \to 0$.

En contournant le point z_j dans le sens négatif la fonction $\ln f(z) = \ln |f(z)| + i \arg f(z)$ augmente de $-2\pi k_j i$; par conséquent, l'intégrale le long de la partie du contour composée par le segment L_j parcouru deux fois dans les sens opposés vaut

$$k_j \int_{L_j} \frac{dz}{z} = k_j [\ln z_j' - \ln z_j''].$$

Sur chaque partie suivante de la circonférence $|z| = r$, la fonction $\ln f(z)$ augmente de $-2\pi k_j i$, ce qui apporte à l'intégrale (1) la quantité

$$k_j \int_{\varphi_j}^{\varphi_{j+1}} i\, d\theta$$

qui est évidemment purement imaginaire. Ceci étant, séparons la partie réelle dans l'égalité (1) pour $\varepsilon \to 0$; nous obtenons:

$$\ln|f(0)| = \sum_{j=1}^{m} k_j \ln|z_j| + \frac{1}{2\pi}\int_0^{2\pi} \ln|f(re^{i\theta})|\, d\theta.$$

Or, puisque $|z_j| < 1$, $\ln|z_j| < 0$, on a

$$\frac{1}{2\pi}\int_0^{2\pi} \ln|f(re^{i\theta})|\, d\theta \geqslant \ln|f(0)|$$

ou bien, ce qui revient au même

$$-\frac{1}{2\pi}\int_0^{2\pi} \ln|f(re^{i\theta})|\, d\theta \leqslant -\ln|f(0)|$$

Par hypothèse nous avons sur la circonférence $|z| = 1$ un seul zéro au point $z^* = 1$. Choisissons un $\delta > 0$ quelconque; il est évident que

$$-\frac{1}{2\pi}\int_{\delta}^{2\pi-\delta} \ln|f(re^{i\theta})|\, d\theta \leqslant -\frac{1}{2\pi}\int_0^{2\pi} \ln|f(re^{i\theta})|\, d\theta \leqslant -\ln|f(0)|.$$

En gardant δ fixe passons à la limite, dans cette inégalité, pour $r \to 1$. Nous avons

$$-\frac{1}{2\pi}\int_{\delta}^{2\pi-\delta} \ln|f(e^{i\theta})|\, d\theta \leqslant -\ln|f(0)|.$$

Cette inégalité est valable pour tout $\delta > 0$. En passant à la limite pour $\delta \to 0$ nous voyons que l'intégrale

$$-\frac{1}{2\pi}\int_0^{2\pi} \ln|f(e^{i\theta})|\, d\theta$$

existe. Le théorème est démontré

Exercices

1. Démontrer la convergence de l'intégrale de Fourier

$$\varphi(x) = \lim_{p\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} \varphi(x+t) \frac{\sin pt}{t}\, dt$$

en un point x au voisinage duquel la fonction $\varphi(x)$ est continue et monotone, ainsi que sa convergence uniforme dans tout intervalle fermé intérieur à un intervalle de monotonie et de continuité de $\varphi(x)$.

2. Donner un exemple de fonction $\varphi(x)$ continue, possédant une intégrale de Fourier uniformément convergente et telle que la fonction

$$\psi(\sigma) = \int_{-\infty}^{\infty} \varphi(x)\, e^{ix\sigma}\, dx$$

n'est pas absolument intégrable sur $-\infty < \sigma < \infty$.

3. Donner un exemple de fonction $\varphi(x)$ pour laquelle l'intégrale de Fourier

$$\lim_{p\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} \varphi(x+t) \frac{\sin pt}{t}\, dt = \lim_{p\to\infty} \frac{1}{2\pi} \int_{-p}^{p} \int_{-\infty}^{\infty} \varphi(\xi)\, e^{i\sigma(x-\xi)}\, d\xi$$

converge uniformément sur l'axe $-\infty < x < \infty$, mais chacune des intégrales

$$\lim_{p\to\infty} \int_{0}^{p} \left\{ \int_{-\infty}^{\infty} \varphi(\xi)\, e^{i\sigma(x-\xi)}\, d\xi \right\}, \quad \lim_{p\to\infty} \int_{-p}^{0} \left\{ \int_{-\infty}^{\infty} \varphi(\xi)\, e^{i\sigma(x-\xi)}\, d\xi \right\}$$

possède des points de divergence.

4. Démontrer l'« égalité de Parseval »

$$\int_{-\infty}^{\infty} |g(\sigma)|^2\, d\sigma = 2\pi \int_{-\infty}^{\infty} |f(x)|^2\, dx,$$

où $g(\sigma)$ est la transformée de Fourier d'une fonction $f(x)$ vérifiant l'hypothèse du théorème 15.13 et de carré intégrable sur tout l'axe $-\infty < x < \infty$ (Plancherel).

5. Démontrer la « relation d'incertitude »

$$\int_{-\infty}^{\infty} x^2 |f(x)|^2\, dx \cdot \int_{-\infty}^{\infty} \sigma^2 |g(\sigma)|^2\, d\sigma \geqslant \frac{\pi}{2}$$

en supposant que les fonctions $xf(x)$ et $f'(x)$ vérifient l'hypothèse de l'exercice 4 et que $\int_{-\infty}^{\infty} |f(x)|^2\, dx = 1$.

6. Soit $F(p)$ la transformée de Laplace d'une fonction $f(t)$. Trouver les transformées de Laplace des fonctions $f_1(t) = e^{at} f(t)$, $f_2(t) = f'(t)$.

$$f_3(t) = \int_0^t f(\tau)\, d\tau, \quad f_4(t) = tf(t), \quad f_5(t) = \frac{f(t)}{t}.$$

7. Trouver les transformées de Laplace des fonctions

$$\varphi_1(t)=e^{at}\ ;\quad \varphi_2(t)=t^{\alpha-1}\ (\alpha>0)\ ;\quad \varphi_3(t)=t^{\alpha-1}e^{at}\ ;$$

$$\varphi_4(t)=\sin at\ ;\quad \varphi_5(t)=\cos at\ ;\quad \varphi_6(t)=\frac{\sin at}{t}\ .$$

8. Démontrer la formule d'inversion de Mellin : si

$$F(s)=\int_0^\infty f(x)\,x^{s-1}\,dx,$$

alors

$$F(x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F(s)\,x^{-s}\,ds.$$

Historique

L'intégrale de Fourier apparaît pour la première fois dans le livre de Fourier « Théorie analytique de la chaleur » (1822) où elle est appliqué aux plusieurs problèmes de la physique mathématique. Les travaux de Fourier, de même que ceux de Cauchy qui utilise l'intégrale de Fourier en étudiant la propagation des ondes (1842), ne contiennent aucune démonstration de la convergence ; les démonstrations rigoureuses, dans de diverses hypothèses, apparaissent durant tout le XIXe siècle et représentent des modifications des démonstrations correspondantes de la convergence des séries de Fourier. La « transformation de Laplace » est développée par Laplace en 1812 dans la « Théorie analytique des probabilités » ; d'ailleurs Euler considérait dès 1737 les intégrales de $e^{-px}f(x)$ pour résoudre des équations différentielles ordinaires. Au temps d'Euler et Laplace, il n'est point question d'utiliser la transformation de Laplace dans le domaine complexe. A partir do 1892 paraissent les travaux de l'ingénieur anglais Heaviside qui trouve, en interprétant d'après les règles introduites par lui-même les fonctions de $p=\frac{\partial}{\partial t}$ (en dehors de la classe des fonctions rationnelles), les solutions de certains problèmes électrotechniques qui se réduisent à des équations aux dérivées partielles. Un certain temps, le « calcul opérationnel » de Heaviside reste sans fondement mathématique. A compter de 1910, Bromwich, puis Carson, Van der Pol, Doetsch, en appliquant la transformation de Laplace dans le domaine complexe, justifient les règles de Heaviside. Les travaux de Denjoy, Carleman, Ostrovski sur les classes quasi analytiques de fonctions se rapportent aux années 1920-1930.

Le progrès ultérieur dans la théorie de la transformation de Fourier est lié, d'une part, avec l'utilisation de l'intégrale de Lebesgue (et de celle de Lebesgue-Stieltjes) et, d'autre part, avec la théorie des distributions ; en particulier, les distributions permettent de définir la transformée de Fourier pour une fonction indéfiniment croissante (pour $|x|\to\infty$). A son tour, ce fait se trouve essentiel pour la résolution des problèmes fondamentaux de la théorie des équations linéaires aux dérivées partielles et à coefficients constants. Cf. [13], [15] et [10].

Courbes gauches

> La mathématique pure a pour objet les formes spatiales et les rapports quantitatifs du monde réel, donc une matière très concrète. Que cette matière apparaisse sous une forme extrêmement abstraite, ce fait ne peut masquer que d'un voile superficiel son origine située dans le monde extérieur.
>
> *F. Engels*

§ 16.1. Définitions de base

16.11. Par définition, une *courbe* L dans un espace n-dimensionnel R_n est un lieu géométrique caractérisé par un système d'équations paramétriques

$$x_1 = x_1(t), \ \ldots, \ x_n = x_n(t) \quad (a \leqslant t \leqslant b), \tag{1}$$

ou, ce qui revient au même, par une équation vectorielle

$$x = x(t) \quad (a \leqslant t \leqslant b). \tag{2}$$

Le vecteur $x(t)$ s'appelle *rayon vecteur* de la courbe L.

Les fonctions $x_j(t)$ sont supposées continues et vérifiant certaines conditions de dérivabilité qui seront précisées dans la suite. Pour qu'un lieu géométrique (1) ait la forme habituelle d'une courbe, il ne suffit pas que les fonctions $x_j(t)$ soient continues: il existe des systèmes (1), avec les deuxièmes membres continus, pour lesquels le lieu géométrique correspondant représente tout l'espace R_n (cf. exercice 5). Remarquons encore qu'une même courbe (i.e. un même lieu géométrique) peut être donnée par plusieurs systèmes différents (1); par exemple, pour $-\infty < t < \infty$, les systèmes

$$x = r\cos t, \quad y = r\sin t$$

et

$$x = r\cos(t^3), \quad y = r\sin(t^3)$$

définissent un même lieu géométrique dans le plan (x, y), à savoir la circonférence de rayon r centrée à l'origine des coordonnées. Nous verrons dans ce qui suit que le choix de la représentation paramétrique convenable aide souvent à expliciter les propriétés géométriques d'une courbe donnée.

16.12. Nous avons déjà vu un cas plus général où $x(t)$ de l'équation 16.11(2) désignait un point d'un espace métrique dépendant d'un paramètre t; c'était *une courbe dans un espace métrique*. Dans 12.61, pour le cas où les valeurs de la fonction $x(t)$ appartiennent à un espace normé B, nous avons défini la dérivée de la fonction vectorielle $x(t)$ en un point $t = c$:

$$x'(c) = \lim_{\Delta t \to 0} \frac{x(c+\Delta t) - x(c)}{\Delta t} \tag{1}$$

si la limite dans le deuxième membre existe au sens de la métrique de l'espace B. Alors la fonction $x(t)$ est *dérivable au point* $t = c$. Si la limite (1) existe pour tout $c \in [a, b]$, alors la fonction $x(t)$ est *dérivable sur l'intervalle* $[a, b]$.

Nous allons étudier les fonctions dérivables à valeurs dans un espace m-dimensionnel, mais certains résultats seront valables pour un espace normé (de dimension infinie).

16.13. Remarquons avant tout que puisque le passage à la limite dans un espace m-dimensionnel est équivalent au passage à la limite pour chaque coordonnée, la dérivabilité pour $t = c$ de la fonction vectorielle $x(t) = (x_1(t), \ldots, x_m(t))$ au sens de 16.12(1) est équivalente à la dérivabilité de m fonctions numériques $x_1(t), \ldots, x_m(t)$ pour $t = c$; de plus, on a

$$x'(c) = (x_1'(c), \ldots, x_m'(c)) \in R_m. \tag{1}$$

Interprétons géométriquement la dérivabilité d'une fonction vectorielle. On en a déjà parlé au début du chapitre 13; ici nous traitons la question indépendamment et plus en détail.

La définition 16.12(1) de la dérivée peut être mise sous une forme équivalente

$$\Delta x = x(c + \Delta t) - x(c) = x'(c)\,\Delta t + \varepsilon(t)\,\Delta t, \tag{2}$$

où le vecteur $\varepsilon(t)$ tend vers zéro pour $\Delta t \to 0$. L'égalité (2) montre que l'accroissement de la fonction $x(t)$ lorsque t varie de c à $c + \Delta t$ contient la partie linéaire principale $x'(c)\,\Delta t$. Un point avec $t = c$ sera dit *ordinaire* si $x'(c) \neq 0$ et *singulier* si $x'(c) = 0$ (cf. 9.63*g*). Pour un point ordinaire, l'image géométrique correspondant à l'équation linéaire $z = x(c) + x'(c)(t - c)$ est la droite Λ passant par le point $M = x(c)$ dans le sens du vecteur $x'(c)$ (fig. 16.1). Ainsi, l'écart d'un point de la courbe du point correspondant (i.e. pour la même valeur de t) de la droite Λ est un infiniment petit d'ordre supérieur par rapport à Δt. Pour cette raison, la droite Λ s'appelle *tangente à la courbe L au point M*. De la sorte, *l'existence d'une*

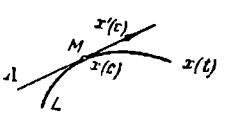


Fig. 16.1.

dérivée $x'(c)$ non nulle est équivalente à l'existence d'une tangente à la courbe L au point M ; le vecteur $x'(c)$ est un vecteur directeur de cette tangente.

16.14. Voyons ce qui arrive à un vecteur directeur d'une tangente lorsqu'on passe, sur la courbe L, à un nouveau paramètre τ, de sorte que $t = t(\tau)$ soit une fonction dérivable de τ. Soient, en particulier, $c = t(\gamma)$ et $t'(\gamma) \neq 0$. Alors nous posons $x(t(\tau)) \equiv g(\tau)$ et écrivons

$$g'(\tau) = \lim_{\Delta\tau\to 0} \frac{\Delta x}{\Delta \tau} = \lim_{\Delta t\to 0} \frac{\Delta x}{\Delta t} \lim_{\Delta\tau\to 0} \frac{\Delta t}{\Delta \tau} = x'(t)\, t'(\tau) \qquad (1)$$

d'après la règle de dérivation d'une fonction composée (12.61*h*).

Donc, le nouveau vecteur directeur $g'(\tau)\,|_{\tau=\gamma}$ a le même sens que l'ancien vecteur $x'(t)$ et s'en déduit par la multiplication par $t'(\tau)$. Ainsi, la longueur du vecteur directeur d'une tangente n'admet aucune interprétation géométrique directe. Comme nous l'avons vu dans 13.11*d*, on peut attribuer au vecteur $x'(t)$ un sens cinématique ; si t est le temps, alors $x'(c)$ est la vitesse du mouvement du point $x = x(t)$ le long de la courbe L à l'instant $t = c$.

16.15. Introduisons, enfin, la notion de différentielle d'une fonction vectorielle $x(t)$. Le vecteur $dx = x'(c)\, dt$, où $dt = \Delta t$ est un accroissement arbitraire du paramètre t, s'appelle *différentielle de la fonction vectorielle $x(t)$* pour $t = c$. La différentielle d'une fonction est donc la partie linéaire principale de son accroissement correspondant à un accroissement de la variable indépendante t.

Comme plus haut, on a le théorème sur l'invariance de la différentielle : *la différentielle d'une fonction a la même forme que t soit une variable indépendante ou une fonction d'une autre variable indépendante* (dans le dernier cas, dt est la partie linéaire principale de l'accroissement de la fonction $t(\tau)$). En effet, si $g(\tau) = x[t(\tau)]$, alors

$$d_{(\tau)}x = g'(\gamma)\, d\tau = x'(c)\, t'(\gamma)\, d\tau = x'(c)\, dt = d_{(t)}x,$$

ce qu'on affirmait.

16.16. Intégration d'une fonction vectorielle. Dans 12.62 nous avons défini cette opération pour les fonctions vectorielles à valeurs dans un espace normé B. On appelle intégrale d'une fonction vectorielle $x(t)$, $a \leqslant t \leqslant b$, à valeurs dans un espace complet B, dans l'espace R_m par exemple, la quantité

$$\int_a^b x(t)\, dt = \lim_{d(\Pi)\to 0} \sum_{k=0}^{n-1} x(\xi_k)\, \Delta t_k,$$

$$\Pi = \{a = t_0 \leqslant \xi_0 \leqslant t_1 \leqslant \ldots \leqslant t_n = b\}, \quad \Delta t_k = t_{k+1} - t_k,$$

où il s'agit de la limite par rapport à la norme de l'espace B pour le morcellement illimité de la partition Π, i.e. pour

$$d(\Pi) = \max \Delta t_k \to 0.$$

L'existence de l'intégrale a été démontrée pour $x(t)$ continue par morceaux. Les propriétés principales de l'intégrale sont indiquées dans 12.62*c* :

1) $\int_a^b \alpha x(t)\,dt = \alpha \int_a^b x(t)\,dt$ pour tout α réel ;

2) $\int_a^b [x(t) + y(t)]\,dt = \int_a^b x(t)\,dt + \int_a^b y(t)\,dt$;

3) $\int_a^c x(t)\,dt + \int_c^b x(t)\,dt = \int_a^b x(t)\,dt$ $(a \leqslant c \leqslant b)$;

4) $\left\| \int_a^b x(t)\,dt \right\| \leqslant \max_{a \leqslant t \leqslant b} \| x(t) \| (b-a)$.

Complétons-les par la formule d'intégration par parties

5) $\int_a^b u(t)\,dv(t) = u(t)\,v(t)\Big|_a^b - \int_a^b v(t)\,du(t)$.

Ici $u(t)$ est une fonction dérivable à valeurs dans l'espace B, $v(t)$ une fonction numérique dérivable. La démonstration de la formule 5) est analogue à celle pour les fonctions numériques (9.51*a*).

16.17. Dérivés d'ordre supérieur.

a. Les dérivées d'ordre supérieur d'une fonction vectorielle $x(t)$ sont définies dans 12.64. La *dérivée* n-ième est, par définition, la dérivée première de la dérivée $(n-1)$-ième si cette dernière est une fonction dérivable pour $a \leqslant t \leqslant b$.

Dans ce qui suit, l'existence de ces dérivées sera toujours supposée.

b. Voyons comment changent les fonctions vectorielles x_t, x_{tt}, ... lorsqu'on remplace la variable indépendante t par une nouvelle variable indépendante τ, $t = t(\tau)$, où $t(\tau)$ est une fonction suffisamment lisse de τ.

Nous avons déjà vu (16.14) que les dérivées premières en t et en τ ne diffèrent que par le facteur

$$x_\tau = x_t t_\tau.$$

En dérivant cette égalité en τ et en appliquant encore une fois la formule de dérivation d'une fonction composée on trouve

$$x_{\tau\tau} = (x_\tau)_\tau = (x_t t_\tau)_\tau = (x_t)_\tau t_\tau + x_t t_{\tau\tau} = x_{tt} t_\tau^2 + x_t t_{\tau\tau}.$$

Nous voyons que le vecteur $x_{\tau\tau}$, sans être en général colinéaire au vecteur x_{tt}, se trouve dans le plan des vecteurs x_t et x_{tt}. Ainsi, le plan déterminé par les vecteurs x_t et x_{tt} ne dépend pas du choix du paramètre, bien que la position du vecteur x_{tt} dans le plan change lorsqu'on passe à un nouveau paramètre.

Dans le cas général, *quel que soit n, le vecteur* $x_\tau^{(n)}$ *appartient au sous-espace n-dimensionnel engendré par les vecteurs* $x_t, x_{tt}, \ldots, x_t^{(n)}$. On le démontre par récurrence : en supposant la relation

$$x_\tau^{(n)} = \sum_{k=1}^{n} x_t^{(k)} \varphi_k(\tau)$$

on la dérive encore une fois par rapport à τ ; on obtient

$$x_\tau^{(n+1)} = \sum_{k=1}^{n} x_t^{(k+1)} t_\tau \varphi_k(\tau) + \sum_{k=1}^{n} x_t^{(k)} \varphi_k'(\tau),$$

de sorte que $x_\tau^{(n+1)}$ s'exprime linéairement par $x_t, x_{tt}, \ldots, x_t^{(n+1)}$, ce qu'il nous fallait.

On peut dire que ce ne sont pas les vecteurs $x_t, x_{tt}, \ldots, x_t^{(n)}, \ldots$ eux-mêmes qui ont un sens géométrique, mais les variétés linéaires qu'ils engendrent. Ces variétés linéaires sont appelées *sous-espaces osculateurs* de dimension 1, 2, ..., n, ... (dans le cas où x_t, x_{tt}, ... sont linéairement indépendants).

c. Si la fonction $x(t)$ est $n+1$ fois dérivable sur l'intervalle $[a, b]$, alors la formule de Taylor (12.64 *c*) est valable :

$$\Delta x(t) = x(t+\Delta t) - x(t) =$$
$$= x'(t)\Delta t + x''(t)\frac{(\Delta t)^2}{2!} + \ldots + x^{(n)}(t)\frac{(\Delta t)^n}{n!} + Q_n.$$

En y posant $n = 1, 2, \ldots$ et en utilisant l'estimation pour Q_n de 12.64*c* nous aboutissons à une série de formules de plus en plus précises :

$$\Delta x(t) = x'(t)\Delta t + \varepsilon_1(t)\Delta t, \tag{1}$$

$$\Delta x(t) = x'(t)\Delta t + x''(t)\frac{(\Delta t)^2}{2} + \varepsilon_2(t)(\Delta t)^2, \tag{2}$$

$$\Delta x(t) = x'(t)\Delta t + x''(t)\frac{(\Delta t)^2}{2} + x'''(t)\frac{(\Delta t)^3}{6} + \varepsilon_3(t)(\Delta t)^3, \tag{3}$$

. . . . . .

16.18. Forme d'une courbe au voisinage d'un point ordinaire ou singulier. L'égalité 16.17(1) montre qu'une courbe L d'équation $x = x(t)$ coïncide, à un infini-

ment petit du premier ordre par rapport à Δt près, avec sa tangente si $x'(t) \neq 0$. L'égalité 16.17(2) montre que la courbe L se trouve, à un infiniment petit du deuxième ordre près, dans le plan défini par les vecteurs $x'(t)$ et $x''(t)$; conformément à la définition 16.17*b*, ce plan (dans le cas où $x'(t)$ et $x''(t)$ sont linéairement indépendants) s'appelle *plan osculateur à la courbe au point M*. Si ξ, η sont les coordonnées dans le plan osculateur par rapport à la base $x'(t)$, $x''(t)/2$, alors nous obtenons à partir de 16.17(2) la représentation paramétrique de la courbe L à un infiniment petit du deuxième ordre près :

$$\xi = \Delta t, \quad \eta = (\Delta t)^2.$$

Par conséquent, avec la précision indiquée, la courbe L est une parabole dans le plan osculateur avec $\eta = \xi^2$ pour équation.

Le sous-espace défini par les vecteurs $x'(t)$, $x''(t)/2$, $x'''(t)/6$ (en cas de leur indépendance linéaire) s'appelle *sous-espace osculateur tridimensionnel à la courbe L au point M* (16.17*b*). On voit dans 16.17(3) que la courbe L se trouve, à un infiniment petit du troisième ordre près, dans son sous-espace osculateur tridimensionnel. Si ξ, η, ζ sont les coordonnées dans ce sous-espace par rapport à la base $x'(t)$, $x''(t)/2$, $x'''(t)/6$, alors on déduit de la même égalité 16.17(3) la représentation paramétrique de la courbe L à un infiniment petit du troisième ordre près :

$$\xi = \Delta t, \quad \eta = (\Delta t)^2, \quad \zeta = (\Delta t)^3.$$

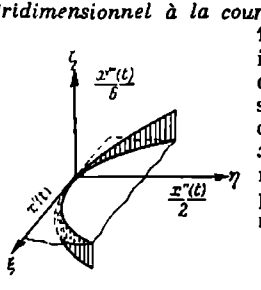


Fig. 16.2.

Nous obtenons une courbe gauche montrée fig. 16.2. Sa projection sur le plan des ξ, η est la parabole déjà considérée $\eta = \xi^2$. Sa projection sur le plan des ξ, ζ est la courbe du troisième degré $\zeta = \xi^3$ (vue à partir de l'extrémité du vecteur $\frac{1}{2}x''(t)$). Sa projection sur le plan des η, ζ est la parabole semi-cubique $\zeta = \eta^{3/2}$ (vue à partir de l'extrémité du vecteur $x'(t)$).

Considérons à présent la courbe L au voisinage d'un point singulier c, où $x'(c) = 0$ mais $x''(c) \neq 0$ et $x'''(c) \neq 0$. La formule de Taylor nous donne

$$\Delta x(c) = \frac{1}{2}x''(c)\,\Delta t^2 + \frac{1}{6}x'''(c)\,\Delta t^3 + \varepsilon_3(t)\,\Delta t^4.$$

Ainsi, à un infiniment petit du troisième ordre près, la courbe L se trouve dans le plan défini par les vecteurs $x''(c)/2$, $x'''(c)/6$ et a dans ce plan l'équation $\zeta = \eta^{3/2}$. Si l'on garde les infiniment petits du quatrième ordre, cela fournit le terme complémentaire

$\frac{1}{24} x^{IV}(c)\Delta t^4$ (pour $x^{IV}(c) \neq 0$) qui montre que la courbe s'écarte du plan $x''(c)$, $x'''(c)$ dans le demi-espace indiqué par le vecteur $x^{IV}(c)$ (fig. 16.3). Remarquons que, le signe de Δt^4 étant constant, les deux branches de la pointe s'écartent du plan dans un même demi-espace.

Ainsi, en cas de singularité c avec $x'(c) = 0$, $x''(c) \neq 0$, $x'''(c) \neq 0$, *le point singulier M est un point de rebroussement.*

16.19. Longueur d'arc. La définition de la longueur d'arc d'une courbe $x(t)$ est donnée dans 9.63. Nous allons reproduire cette définition et déduire la formule correspondante relativement à la courbe dans un espace normé quelconque. La longueur d'un arc de courbe est définie comme limite des longueurs des lignes polygonales inscrites lorsque la longueur de chaque segment diminue indéfiniment. D'une façon plus exacte, soit

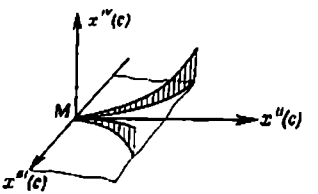


Fig. 16.3.

$$\Pi = \{a = t_0 < t_1 < \ldots < t_n = b\}$$

une partition de l'intervalle $[a, b]$ sur lequel est définie une fonction vectorielle $x(t)$. A tout point t_i il correspond sur la courbe un point $M_i = x(t_i)$. En joignant les points M_i par les segments de droite nous avons une ligne polygonale L_Π dont la longueur vaut $\sum_{i=0}^{n-1} |\Delta x_i|$. Supposons que la fonction $x(t)$ soit continûment dérivable sur l'intervalle $[a, b]$. Alors

$$\Delta x_i = \int_{t_i}^{t_{i+1}} x'(t)\,dt = x'(t_i)\,\Delta t_i + \varepsilon_i \Delta t_i,$$

où

$$\varepsilon_i = \frac{1}{\Delta t_i} \int_{t_i}^{t_{i+1}} |x'(t) - x'(t_i)|\,dt,$$

$$|\varepsilon_i| \leqslant \varepsilon_\Pi = \max_{|t-\bar{t}| \leqslant d(\Pi)} |x'(t) - x'(\bar{t})|.$$

Vu la continuité uniforme de la fonction $x'(t)$, cette quantité tend vers zéro pour un morcellement illimité de la partition Π. Nous avons donc l'estimation

$$\Big|\sum_{i=0}^{n-1} |\Delta x_i| - \sum_{i=0}^{n-1} |x'(t_i)|\,\Delta t_i\Big| \leqslant \sum_{i=0}^{n-1} |\varepsilon_i|\,\Delta t_i \leqslant \varepsilon_\Pi (b-a).$$

Or, la somme $\sum_{i=0}^{n-1} |x'(t_i)| \Delta t_i$ tend, pour un morcellement illimité de la partition Π, vers la limite

$$\int_a^b |x'(t)|\,dt, \tag{1}$$

car la fonction numérique $|x'(t)|$ est continue dès que la fonction vectorielle $x'(t)$ l'est. Il en résulte que la limite des longueurs des lignes polygonales inscrites existe et vaut l'intégrale (1). Notons que, dans l'espace R_n, nous avons $|x'(t)| = \sqrt{\sum_{k=1}^{n} [x_k'(t)]^2}$, ce qui correspond à la formule déjà vue 9.63(5). Contrairement à cette formule, l'expression (1) est valable pour tout espace normé.

Si l'on remplace b par t et t par τ, on obtient l'expression pour la longueur d'arc de la courbe L relative à l'intervalle $[a, t]$ de variation du paramètre τ :

$$s(t) = \int_a^t |x'(\tau)|\,d\tau.$$

Nous voyons que $s(t)$ est une fonction non décroissante de t, continue et dérivable ; de plus, on a

$$s'(t) = |x'(t)|$$

d'après 9.31.

Si la courbe L n'a pas de points singuliers, i.e. $x'(t)$ ne s'annule en aucun point, on peut appliquer le théorème sur la fonction inverse ; il existe alors une fonction inverse $t = t(s)$ qui est continue, croissante et continûment dérivable. Ceci étant, on peut mettre la fonction $x(t)$ sous la forme d'une fonction de s, toujours continue et continûment dérivable. La longueur d'arc s sera appelée *paramètre naturel*. Si la courbe L est donnée par une fonction $x = x(s)$ avec le paramètre naturel s, alors

$$|x'(s)| = s'(s) = 1$$

en vertu de (1).

Ainsi, *en tout point non singulier de la courbe L le vecteur* $x'(s)$ *a 1 pour longueur.* (Ceci est clair du point de vue cinématique : si le paramètre s est à la fois le chemin parcouru et le temps employé, alors la vitesse du mouvement est égale à l'unité.)

§ 16.2. Courbure. Courbures d'ordre supérieur

16.21. Dans ce qui suit, nous aurons affaire non seulement aux longueurs de vecteurs, mais aussi aux angles qu'ils forment. Il est donc naturel d'opérer non pas dans un espace normé quelconque mais dans un espace hilbertien (12.41).

Lemme. *Soient $x(t)$ et $y(t)$ $(a \leqslant t \leqslant b)$ deux fonctions dérivables à valeurs dans un espace hilbertien* H; *alors la fonction numérique* $\varphi(t) = (x(t), y(t))$ *est dérivable elle aussi, et l'on a*

$$\varphi'(t) = (x'(t), y(t)) + (x(t), y'(t)). \tag{1}$$

Démonstration. Nous avons

$$\begin{aligned}\Delta\varphi &= (x(t+\Delta t), y(t+\Delta t)) - (x(t), y(t)) = \\ &= (x(t) + x'(t)\,\Delta t + \varepsilon_1\Delta t,\ y(t) + \\ &+ y'(t)\,\Delta t + \varepsilon_2\Delta t) - (x(t), y(t)) = \\ &= [(x'(t), y(t)) + (x(t), y'(t))]\,\Delta t + \varepsilon_3\Delta t,\end{aligned}$$

où $\varepsilon_3 = \varepsilon_3(t, \Delta t) \to 0$ pour $\Delta t \to 0$. On peut donc dégager la partie linéaire principale de l'accroissement $\Delta\varphi$. En l'explicitant on arrive à la formule (1).

Conséquence. *Si la longueur du vecteur $x(t)$ reste constante lorsque t varie, alors le vecteur $x'(t)$ est orthogonal à $x(t)$.*

En effet, en appliquant à la fonction $(x(t), x(t))$ la formule de dérivation (1), nous avons

$$0 = (x(t), x(t))' = 2(x(t), x'(t)),$$

ce qu'il nous fallait.

16.22. Considérons une courbe $L = \{x = x(s)\}$ avec pour paramètre la longueur d'arc comptée à partir d'un point fixe. Comme nous l'avons vu dans 16.19, le vecteur $e_1(s) = x'(s)$ est unitaire.

Si les vecteurs $x'(s)$ et $x''(s)$ sont linéairement indépendants, il existe un plan osculateur. Le vecteur $e_1'(s) = x''(s)$ appartenant au plan osculateur est, on l'a vu, orthogonal à $e_1(s)$. Il s'appelle *vecteur courbure* de la courbe L au point s. Posons

$$e_1'(s) = \varkappa(s)\, e_2(s), \tag{1}$$

où $e_2(s)$ est un vecteur unitaire orthogonal à $e_1(s)$, le coefficient $\varkappa(s)$ étant positif. Nous avons

$$\varkappa(s) = |e_1'(s)| = \lim_{\Delta s \to 0} \left| \frac{e_1(s+\Delta s) - e_1(s)}{\Delta s} \right|. \tag{2}$$

Le module de la différence des vecteurs unitaires $e_1(s + \Delta s)$ et $e_1(s)$ est une corde du cercle unité et représente un infiniment

petit équivalent à l'angle formé par ces vecteurs, i. e. à l'angle des tangentes à la courbe L aux points correspondant respectivement aux valeurs s et $s + \Delta s$. Ainsi, le coefficient $\varkappa(s)$ donne la vitesse de rotation de la tangente par rapport à la variation de la longueur d'arc. Le nombre $\varkappa(s)$ s'appelle *courbure* de la courbe L au point s.

Notons que la formule (2) représente une définition plus générale que la formule (1) car (2) n'exige pas la condition $e_1'(s) \neq 0$, il suffit que $e_1'(s)$ existe. Dans le cas où $e_1'(s) = 0$, la formule (2) donne pour la courbure au point correspondant la valeur nulle.

16.23. Déduisons la formule pour la courbure dans le cas où la courbe L est donnée par une équation $x = x(t)$ avec un t arbitraire. Comme $s_t = |x_t| = \sqrt{(x_t, x_t)}$ (16.19), on a

$$s_{tt} = \frac{(x_t, x_{tt})}{\sqrt{(x_t, x_t)}} = \frac{(x_t, x_{tt})}{|x_t|}$$

en vertu du lemme 16.21.

Nous avons ensuite

$$x_s = x_t t_s, \quad x_{ss} = x_{tt} t_s^2 + x_t t_{ss} = \frac{x_{tt}}{s_t^2} + x_t \left(\frac{1}{s_t}\right)_t \frac{1}{s_t} =$$

$$= \frac{x_{tt}}{|x_t|^2} - \frac{x_t}{s_t^2} \frac{s_{tt}}{s_t} = \frac{x_{tt}}{|x_t|^2} - \frac{x_t (x_t, x_{tt})}{|x_t|^4}$$

et, définitivement,

$$\varkappa(s) = |x_{ss}| = \left| \frac{x_{tt}}{|x_t|^2} - \frac{x_t (x_t, x_{tt})}{|x_t|^4} \right|. \tag{1}$$

La déduction de cette formule se base sur la définition 16.22(2). Elle est donc valable dans les cas $\varkappa(s) \neq 0$ et $\varkappa(s) = 0$.

En portant la valeur de la courbure dans la formule de Taylor 16.17(2) (avec $e_1'(s) \neq 0$) nous mettons cette dernière sous la forme

$$\Delta x(s) = x'(s)\,\Delta s + x''(s)\frac{\Delta s^2}{2} + \varepsilon_2(s)\,\Delta s^2 =$$

$$= e_1(s)\,\Delta s + \frac{1}{2}\varkappa(s)\,e_2(s)\,\Delta s^2 + \varepsilon_2(s)\,\Delta s^2. \tag{2}$$

16.24. Calculons la courbure de la circonférence. En coordonnées liées avec son plan, l'équation de la circonférence est de la forme

$$x(t) = \{R\cos t,\ R\sin t\}.$$

En dérivant nous avons

$$x_t = \{-R\sin t,\ R\cos t\}, \quad |x_t| = R,$$

$$x_{tt} = \{-R\cos t,\ -R\sin t\}, \quad (x_t, x_{tt}) = 0.$$

Il en découle

$$\varkappa(s) = \frac{|x_{tt}|}{|x_t|^2} = \frac{1}{R},$$

donc *la courbure de la circonférence est l'inverse de son rayon.*

16.25. Soit maintenant $L = \{x = x(s)\}$ une courbe gauche. Si l'on considère dans le plan osculateur de divers cercles tangents à la courbe L, alors l'écart entre les points correspondants de la courbe L et de chaque cercle Q est en général du deuxième ordre de petitesse par rapport à Δs. Essayons, en choisissant convenablement le rayon du cercle tangent, d'obtenir l'écart non pas du deuxième mais du troisième ordre. Soit $z = z(s)$ l'équation d'un cercle tangent, avec le paramètre naturel s et avec le vecteur courbure de même sens que pour L (fig. 16.4). Alors on a d'après la formule 16.23(2):

$$\Delta x(s) = e_1(s)\,\Delta s + \frac{1}{2}\varkappa(s)\,e_2(s)\,\Delta s^2 + \varepsilon_2 \Delta s^2,$$

$$\Delta z(s) = e_1(s)\,\Delta s + \frac{1}{2R}\,e_2(s)\,\Delta s^2 + \bar{\varepsilon}_2 \Delta s^2,$$

où ε_2, $\bar{\varepsilon}_2$ sont des infiniment petits pour $\Delta s \to 0$; notre problème sera donc résolu si l'on pose $R = \frac{1}{\varkappa(s)}$. Le cercle tangent qui se trouve dans le plan osculateur à la courbe L au point s, avec le même sens du vecteur courbure que pour L et de rayon $R = \frac{1}{\varkappa(s)}$, est appelé *cercle osculateur* et son centre *centre de courbure de la courbe L au point s*. Le nombre $R = \frac{1}{\varkappa(s)}$ s'appelle *rayon de courbure de la courbe L au point s*. Si la courbe L est un cercle, son rayon de courbure se confond d'après 16.24 avec son rayon ordinaire.

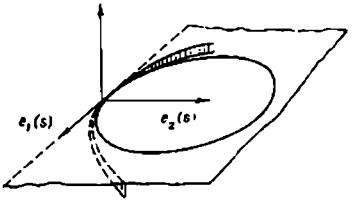


Fig. 16.4.

16.26. Base naturelle. Supposons que, pour un point donné M de la courbe L, les vecteurs $x'(s)$, $x''(s)$, ..., $x^{(n)}(s)$ existent et sont linéairement indépendants. Alors il existe en ce point les sous-espaces osculateurs $E_1 \subset E_2 \subset \ldots \subset E_n$ de dimension 1, 2, ..., n respectivement. Construisons une base orthogonale et normée du sous-espace E_n. Pour deux premiers vecteurs choisissons les vecteurs déjà construits $e_1(s) = x'(s)$ et $e_2(s) = \frac{x''(s)}{|x''(s)|}$. Le troisième vecteur $e_3(s)$ que nous choisissons dans le sous-espace E_3 doit être orthogonal au plan E_2 et de même orientation par rapport à ce plan que le vecteur $x'''(s)$. De façon analogue, une fois choisis les $m-1$ premiers vecteurs $e_1(s)$, ..., $e_{m-1}(s)$, nous construisons dans le sous-espace E_m le vecteur $e_m(s)$ orthogonal au sous-espace E_{m-1} et de même orientation par rapport à ce sous-espace que le vecteur $x^{(m)}(s)$. Ces conditions déterminent d'une façon univoque

la base $e_1(s), \ldots, e_n(s)$. Par construction, chaque vecteur $e_m(s)$ est une combinaison linéaire des vecteurs $x'(s), \ldots, x^{(m)}(s)$:

$$e_m(s) = \varphi_1(s)\, x'(s) + \ldots + \varphi_m(s)\, x^{(m)}(s), \tag{1}$$

avec $\varphi_m(s) > 0$.

La base $e_1(s), \ldots, e_n(s)$ s'appelle *base naturelle de la courbe L* au point M. Il va de soi qu'elle change de position avec le point M.

16.27. Formules de Frénet. Etablissons les formules de dérivation des vecteurs de la base naturelle d'une courbe $L \subset R_n$ par rapport au paramètre s. En dérivant l'égalité 16.26(1) et en utilisant la règle 12.61g nous trouvons

$$e'_m(s) = \sum_{j=1}^{m} \varphi'_j(s)\, x^{(j)}(s) + \sum_{j=1}^{m} \varphi_j(s)\, x^{(j+1)}(s), \tag{1}$$

donc le vecteur $e'_m(s)$, pour $m < n$, appartient au sous-espace E_{m+1}; par conséquent,

$$e'_m(s) = a_{m1}(s)\, e_1(s) + \ldots + a_{mm}(s)\, e_m(s) + a_{m,\,m+1}(s)\, e_{m+1}(s). \tag{2}$$

Pour $m = n$, la formule (2) est valable si l'on remplace e_{n+1} par 0. Il est aisé de déduire de (1), (2) et 16.26(1) que $a_{m,\,m+1} = \varphi_m(s) > 0$. Quant aux autres coefficients dans (2), ils sont aussi faciles à calculer. Tout d'abord nous avons $a_{mm}(s) \equiv 0$ car la dérivée du vecteur unitaire $e_m(s)$ lui est orthogonale (16.21). Ensuite, en dérivant l'égalité évidente $(e_j(s),\ e_m(s)) = 0$ $(j < m)$ on obtient

$$(e'_j(s),\ e_m(s)) + (e_j(s),\ e'_m(s)) = 0,$$

d'où

$$(e'_m(s),\ e_j(s)) = -(e'_j(s),\ e_m(s)); \tag{3}$$

cette expression est nulle pour $j < m - 1$ puisque $e'_j(s) \in E_{j+1}$. Ainsi, la rolation (2) prend la forme

$$e'_m(s) = a_{m,\,m-1}(s)\, e_{m-1}(s) + a_{m,\,m+1}(s)\, e_{m+1}(s). \tag{4}$$

Pour $j = m - 1$, il résulte de (3) que

$$a_{m,\,m-1} = (e'_m(s),\ e_{m-1}(s)) = -(e'_{m-1}(s),\ e_m(s)) = -a_{m-1,\,m}.$$

Posons $\varkappa_1 \equiv \varkappa_1(s) = (e'_1(s),\ e_2(s)) = a_{12}$; dans ce qui précède, cette quantité est désignée par $\varkappa(s)$. Ensuite, posons $\varkappa_2 \equiv \varkappa_2(s) = (e'_2,\ e_3)$, $\varkappa_3 \equiv \varkappa_3(s) = (e'_3,\ e_4)$, etc. Nous aboutissons à un système de formules

$$\left.\begin{aligned}
&e'_1(s) = \varkappa_1 e_2(s),\\
&e'_2(s) = -\varkappa_1 e_1(s) + \varkappa_2 e_3(s),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&e'_{n-1}(s) = -\varkappa_{n-1} e_{n-2}(s) + \varkappa_{n-1}(s)\, e_n(s),\\
&e'_n(s) = -\varkappa_{n-1} e_{n-1}(s)
\end{aligned}\right\} \tag{5}$$

qui s'appellent *formules de Frénet dans* R_n. Les quantités

$$\varkappa_2, \ldots, \varkappa_{n-1}$$

sont appelées *courbures de la courbe L au point M d'ordre* 2, 3, ... respectivement; toutes ces quantités sont positives par construction. La courbure deuxième $\varkappa_2$ s'appelle aussi *torsion* de la courbe au point M.

Dégageons le sens géométrique des coefficients $\varkappa_2, \ldots, \varkappa_{n-1}$. L'égalité

$$e'_m(s) = -\varkappa_{m-1}e_{m-1}(s) + \varkappa_m e_{m+1}(s)$$

montre que la vitesse de rotation du vecteur $e_m(s)$ a deux composantes : la première suivant le vecteur $e_{m-1}(s)$ (elle caractérise la rotation du sous-espace E_m dans lui-même), la seconde suivant le vecteur $e_{m+1}(s)$ (elle correspond à la rotation dans la direction orthogonale à E_m), le coefficient $\varkappa_m$ est la vitesse de cette dernière rotation (angle de rotation rapporté à l'arc parcouru). Ainsi, la torsion est la vitesse de rotation du plan osculateur du vecteur e_2 vers le vecteur e_3.

Le nombre $\varkappa_m$ peut donc être interprété comme vitesse de rotation du sous-espace E_m dans la direction orthogonale à lui-même, de même que la courbure $\varkappa_1$ est géométriquement la vitesse de rotation de la tangente (16.22(2)). Tout comme cette dernière définition, la définition géométrique de $\varkappa_m$ est plus générale que celle basée sur les formules de Frénet et exigeant que l'espace E_{m+1} soit non dégénéré : elle ne demande que la non-dégénérescence de l'espace E_m et l'existence de la dérivée $x^{(m+1)}(s)$. Si $x^{(m+1)}(s) = 0$, l'espace E_m a la vitesse de rotation nulle au point correspondant, et la définition géométrique donne pour $\varkappa_m$ la valeur nulle.

16.28. Calcul de courbures supérieures. En algèbre, on appelle produit mixte de n vecteurs $a_1, \ldots, a_n$ d'un espace euclidien et on le note $[a_1, \ldots, a_n]$ le nombre égal au volume du parallélépipède n-dimensionnel construit sur ces vecteurs; si l'on exprime les vecteurs $a_1, \ldots, a_n$ par leurs coordonnées dans une base orthogonale et normée, alors le nombre $[a_1, \ldots, a_n]$ est égal au déterminant du n-ième ordre dont les colonnes sont formées des coordonnées du vecteur correspondant [14; 8.74].

Calculons le produit mixte des vecteurs $x_s, x_{ss}, \ldots, x_s^{(n)}$. Pour le faire, utilisons les formules

$$\begin{aligned}
x_s &= x_t t_s, \\
x_{ss} &= \ldots + x_{tt} t_s^2, \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
x_s^{(n)} &= \ldots\ldots\ldots\ldots + x_t^{(n)} t_s^n
\end{aligned}$$

exprimant les dérivées en s par les dérivées par rapport à n'importe quel autre paramètre t; les points remplacent les vecteurs qui sont

combinaisons linéaires des vecteurs déjà explicités dans les lignes précédentes. D'après les propriétés des déterminants, on a

$$[x_s, \ldots, x_s^{(n)}] = [x_t, \ldots, x_t^{(n)}]\, t_s^{1+\ldots+n}. \tag{1}$$

D'autre part, nous savons que les formules suivantes sont valables :

$$x_s = e_1,$$
$$x_{ss} = \varkappa_1 e_2,$$
$$x_{sss} = \ldots + (\varkappa_1 e_2)_s = \ldots + \varkappa_1 \varkappa_2 e_3,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_s^{(n)} = \ldots\ldots\ldots\ldots\ldots + \varkappa_1 \ldots \varkappa_{n+1} e_n,$$

où les points remplacent toujours les combinaisons linéaires des vecteurs explicités dans les lignes précédentes. D'une manière analogue, nous avons

$$[x_s, \ldots, x_s^{(n)}] = \varkappa_1^{n-1}\varkappa_2^{n-2} \ldots \varkappa_{n-1}\,[e_1, \ldots, e_n] = \varkappa_1^{n-1}\varkappa_2^{n-2} \ldots \varkappa_{n-1}. \tag{2}$$

En comparant (1) et (2) nous obtenons l'égalité

$$\varkappa_1^{n-1}\varkappa_2^{n-2} \ldots \varkappa_{n-1} = [x_t, \ldots, x_t^{(n)}]\, t_s^{\frac{n(n+1)}{2}} = \frac{[x_t, \ldots, x_t^{(n)}]}{|x_t|^{\frac{n(n+1)}{2}}} \tag{3}$$

qui permet de trouver $\varkappa_{n-1}$ d'après $\varkappa_1, \ldots, \varkappa_{n-2}$. Pour $n = 2$ on a la formule de la courbure

$$\varkappa_1 = \frac{[x_t, x_{tt}]}{|x_t|^3} \tag{4}$$

qui paraît plus simple que celle de 16.23. En réalité, la nouvelle formule (4) n'est plus commode qu'en cas de courbe plane, lorsque la quantité $[x_t, x_{tt}]$ peut être mise sous la forme d'un seul déterminant du deuxième ordre. Rappelons que, dans le cas général, le carré du volume du parallélépipède m-dimensionnel construit sur les vecteurs $x_k = \{x_{k1}, \ldots, x_{kn}\}$ $(k = 1, \ldots, m)$ d'un espace n-dimensionnel est égal à la somme des carrés de tous les déterminants d'ordre m de la matrice des coordonnées des vecteurs x_k [14 ; 8.73].

Pour $n > 2$, en divisant membre à membre la formule (3) par la formule analogue

$$\varkappa_1^{n-2}\varkappa_2^{n-3} \ldots \varkappa_{n-2} = \frac{[x_t, \ldots, x_t^{(n-1)}]}{|x_t|^{\frac{(n-1)n}{2}}}$$

nous obtenons

$$\varkappa_1\varkappa_2 \ldots \varkappa_{n-1} = \frac{[x_t, \ldots, x_t^{(n)}]}{[x_t, \ldots, x_t^{(n-1)}]} \frac{1}{|x_t|^n}. \tag{5}$$

Ensuite, en divisant cette formule membre à membre par la formule analogue

$$\varkappa_1\varkappa_2 \ldots \varkappa_{n-2} = \frac{[x_t, \ldots, x_t^{(n-1)}]}{[x_t, \ldots, x_t^{(n-2)}]} \frac{1}{|x_t|^{n-1}}$$

nous trouvons définitivement

$$\varkappa_{n-1} = \frac{[x_t, \ldots, x_t^{(n)}]\,[x_t, \ldots, x_t^{(n-2)}]}{[x_t, \ldots, x_t^{(n-1)}]^2\,|x_t|}.$$

Géométriquement, cela signifie que la courbure $\varkappa_{n-1}$ est égale, au facteur $|x_t|$ près, au rapport de la hauteur du parallélépipède n-dimensionnel construit sur les vecteurs $x_t, \ldots, x_t^{(n)}$ à la hauteur du parallélépipède $(n-1)$-dimensionnel construit sur les vecteurs $x_t, \ldots, x_t^{(n-1)}$.

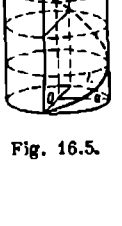


Fig. 16.5.

Exemple. Trouvons la courbure et la torsion d'une hélice dans l'espace tridimensionnel. Une hélice (fig. 16.5) est définie par les équations

$$x(t) = \{a\cos t,\ a\sin t,\ bt\}.$$

En dérivant on obtient

$$x_t = \{-a\sin t,\ a\cos t,\ b\},\quad |x_t| = \sqrt{a^2+b^2},$$
$$x_{tt} = \{-a\cos t,\ -a\sin t,\ 0\},\quad (x_t, x_{tt}) = 0,$$
$$x_{ttt} = \{a\sin t,\ -a\cos t,\ 0\},\quad [x_t, x_{tt}, x_{ttt}] = a^2b.$$

Il en résulte d'après la formule (4):

$$\varkappa_1 = \frac{|x_{tt}|}{|x_t|^2} = \frac{a}{a^2+b^2}.$$

Enfin, vu la formule (3), on a

$$\varkappa_1^2\varkappa_2 = \frac{[x_t, x_{tt}, x_{ttt}]}{|x_t|^6} = \frac{a^2b}{(a^2+b^2)^3},\qquad \varkappa_2 = \frac{b}{a^2+b^2}.$$

§ 16.3. Dégénérescence de la base naturelle

16.31. Dans 16.17, nous avons défini les sous-espaces osculateurs $E_1 \subset E_2 \subset \ldots \subset E_n$ pour la courbe $x = x(t)$ dans le cas où les vecteurs $x'(t), \ldots, x^{(n)}(t)$ existent et sont linéairement indépendants. Considérons le cas où cette dernière condition n'est pas satisfaite.

Si les vecteurs $x'(t), \ldots, x^{(n_1)}(t)$ deviennent linéairement dépendants en un point t (les vecteurs $x'(t), \ldots, x^{(n-1)}(t)$ restant toujours linéairement indépendants), alors l'espace E_n n'existe plus, bien que E_{n-1} continue à exister. La quantité $\varkappa_n(t)$ perd le sens; il en est de même du vecteur $e_n(t)$. On peut essayer de construire le

vecteur $e_n(t)$ par continuité en passant à la limite pour $\bar{t} \to t$ dans $e_n(\bar{t})$ (lorsque $x'(\bar{t}), \ldots, x^{(n)}(\bar{t})$ sont linéairement indépendants), mais il peut s'avérer que les passages à la limite suivant $\bar{t} \nearrow t$ croissants et $\bar{t} \searrow t$ décroissants amènent à des valeurs différentes. Soit, par exemple, C un point d'inflexion d'une courbe plane (fig. 16.6); alors, en passant par ce point, le vecteur $e_2(t)$ se transforme brusquement en son opposé. Nous convenons de compléter la définition des fonctions $\varkappa_1(t), \ldots, \varkappa_n(t)$ en leur attribuant la valeur nulle aux points où elles cessent d'exister; cela est conforme aux considérations de 16.27.

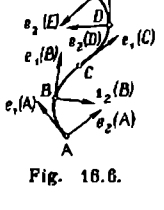


Fig. 16.6.

16.32. Exemple. Une droite est donnée par une équation de la forme

$$x(t) = x_0 + tx_1. \tag{1}$$

Il résulte de (1) que

$$x'(t) = x_1, \quad x''(t) = 0, \tag{2}$$

donc le plan osculateur n'existe pas; par conséquent, la courbure d'une droite est nulle d'après notre convention.

Inversement, supposons que la courbure d'une courbe $x = x(t)$ soit identiquement nulle, i.e. les vecteurs $x'(t)$ et $x''(t)$ soient linéairement dépendants (et $x'(t) \neq 0$). Montrons que c'est une droite. La dépendance linéaire des vecteurs $x'(t)$ et $x''(t)$ veut dire que $x''(s) = 0$ car $x''(s)$ s'exprime linéairement par $x'(t)$ et $x''(t)$ et est orthogonal à $x'(t)$. Alors $x'(s)$ est un vecteur constant (d'après le théorème démontré dans 12.61*k*: si la dérivée d'une fonction vectorielle est identiquement nulle, la fonction est constante) et unitaire, comme tout vecteur de la forme $x'(s)$. Désignons-le par x_1, $|x_1| = 1$. En intégrant l'équation $x'(s) = x_1$ et vu l'unicité de la solution (qui résulte toujours du théorème 12.61*k*) nous avons

$$x(s) = x_0 + sx_1,$$

ce qu'il nous fallait.

16.33. Un cas un peu plus compliqué est une courbe située entièrement dans un hyperplan n-dimensionnel. Cet hyperplan contient tous les vecteurs $x'(t), x''(t), \ldots$, donc les vecteurs $x'(t), \ldots, \ldots, x^{(n+1)}(t)$ sont linéairement dépendants. Par conséquent, la courbure $\varkappa_n(t)$ ainsi que toutes les courbures suivantes sont nulles. Montrons que la réciproque est valable: si la courbure n-ième d'une courbe L est identiquement nulle, alors toute la courbe L se trouve dans un hyperplan n-dimensionnel.

L'hypothèse veut dire que, pour tout t, $a \leqslant t \leqslant b$, les vecteurs $x'(t), \ldots, x^{(n+1)}(t)$ sont linéairement dépendants:

$$x^{(n+1)}(t) = a_0(t)\, x'(t) + \ldots + a_{n-1}(t)\, x^{(n)}(t). \tag{1}$$

Supposons que les vecteurs $x'(t), \ldots, x^{(n)}(t)$ restent linéairement indépendants sur tout l'intervalle $a \leqslant t \leqslant b$. On peut alors montrer que tous les coefficients $a_k(t)$ dans (1) sont continus dès quo la fonction $x^{(n+1)}(t)$ est continue. En effet, en multipliant scalairement l'égalité (1) par $x'(t), \ldots, x^{(n)}(t)$, on obtient le système d'équations linéaires par rapport aux coefficients :

$$(x^{(n+1)}(t),\ x'(t)) = a_0(t)(x_t,\ x_t) + \ldots + a_{n-1}(t)(x_t^{(n)},\ x_t),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$(x^{(n+1)}(t),\ x^{(n)}(t)) = a_0(t)(x_t,\ x_t^{(n)}) + \ldots + a_{n-1}(t)(x_t^{(n)},\ x_t^{(n)}).$$

Tous les coefficients de ce système sont continus (comme produits scalaires de fonctions continues), et son déterminant n'est pas nul comme déterminant de Gramm d'un système de vecteurs linéairement indépendants [14 ; 8.71]. Il résulte des formules de Cramer pour les solutions $a_0(t), \ldots, a_{n-1}(t)$ que ces solutions sont continues sur $[a, b]$.

En vertu de 13.65, il existe une solution $y(t)$ de l'équation

$$y^{(n)}(t) = a_0(t)\,y(t) + \ldots + a_{n-1}(t)\,y^{(n-1)}(t) \qquad (2)$$

qui appartient, d'après 13.66, au sous-espace engendré par les vecteurs $y_0 = y(a), \ldots, y_{n-1} = y^{(n-1)}(a)$. Nous posons ici $y_0 = x'(a), \ldots, y_{n-1} = x^{(n)}(a)$. Comme la fonction vectorielle $x'(t)$ satisfait à l'équation (2) (qui donne (1) après la substitution $x'(t) = y(t)$) et aux conditions initiales indiquées, le vecteur $x'(t)$ reste, d'après 13.66, dans le sous-espace engendré par les vecteurs $x'(a), \ldots, x^{(n)}(a)$ pour tout t. Le vecteur

$$x(t) = x(a) + \int_a^t x'(\xi)\,d\xi$$

appartient, pour tout $t \in [a, b]$, à l'hyperplan passant par le point $x(a)$ parallèlement au sous-espace trouvé. Nous avons démontré le théorème :

Théorème. *Si, pour une courbe $L = \{x = x(t)\}$ dans un espace hilbertien H et pour tout $t \in [a, b]$, les vecteurs $x'(t), \ldots, x^{(n)}(t)$ sont linéairement indépendants et les vecteurs $x'(t), \ldots, x^{(n+1)}(t)$ linéairement dépendants, alors la courbe L appartient à l'hyperplan passant par le point $x(a)$ et défini par les vecteurs $x'(a), \ldots, x^{(n)}(a)$.*

En particulier, si la torsion $\varkappa_2(t)$ de la courbe L est identiquement nulle et la courbure ne s'annule pas, alors la courbe L est plane.

§ 16.4. Equations naturelles

16.41. Etant donnée une courbe L avec le paramètre naturel s, on peut considérer toutes ses courbures comme fonctions de s:

$$\varkappa_1 = \varkappa_1(s),\ \varkappa_2 = \varkappa_2(s),\ \ldots,\quad 0 \leqslant s \leqslant s_0.$$

Nous admettons que les courbures s'annulent en sous-entendant la dégénérescence décrite dans 16.31.

Si une courbe $\bar{L}$ se déduit de la courbe L par une transformation linéaire isométrique (i.e. conservant toutes les distances) de l'espace H, alors, toutes les fonctions $\varkappa_m(s)$, $m = 1, 2, \ldots$ étant complètement déterminées par la métrique, elles resteront les mêmes pour la courbe $\bar{L}$. Montrons que la proposition inverse est valable pour les courbes de dimension finie:

Théorème. *Si, pour deux courbes $\bar{L}$ et L dans l'espace n-dimensionnel R_n représentées par des fonctions vectorielles n fois dérivables, les courbures $\varkappa_1(s), \ldots, \varkappa_{n-1}(s)$ sont continues, positives et s'expriment par les fonctions identiques du paramètre naturel s, alors il existe une transformation isométrique (un déplacement complété éventuellement par une symétrie) de l'espace R_n dans lui-même qui transforme la courbe L en la courbe L.*

Démonstration. Soient $e_1(s), \ldots, e_n(s)$ la base naturelle de la courbe L et $\bar{e}_1(s), \ldots, \bar{e}_n(s)$ celle de $\bar{L}$. Considérons un déplacement (suivi peut-être d'une symétrie) de l'espace R_n qui transforme la position initiale $\bar{e}_1(0), \ldots, \bar{e}_n(0)$ de la base naturelle de $\bar{L}$ en la position initiale $e_1(0), \ldots, e_n(0)$ de la base naturelle de L, de sorte que $\bar{e}_1(0)$ devient $e_1(0), \ldots, \bar{e}_n(0)$ devient $e_n(0)$. Montrons que cette transformation de l'espace R_n transforme $\bar{L}$ en L.

Les fonctions $\varkappa_1(s), \ldots, \varkappa_{n-1}(s)$ sont continues par hypothèse. Il existe donc, vu le théorème 13.63, la solution unique du système

$$\left.\begin{aligned} y_1'(s) &= \varkappa_1(s)\, y_2(s), \\ y_2'(s) &= -\varkappa_1(s)\, y_1(s) + \varkappa_2(s)\, y_3(s), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n'(s) &= -\varkappa_{n-1}(s)\, y_{n-1}(s) \end{aligned}\right\} \tag{1}$$

avec les conditions initiales

$$y_1(0) = e_1(0), \ldots, y_n(0) = e_n(0). \tag{2}$$

Or, d'après les formules de Frénet 16.27(5), les vecteurs $e_1(s), \ldots, e_n(s)$ ainsi que les vecteurs $\bar{e}_1(s), \ldots, \bar{e}_n(s)$ (après déplacement) satisfont au système (1) avec les conditions initiales (2);

par conséquent,

$$e_1(s) \equiv \bar{e}_1(s), \ \ldots, e_n(s) \equiv \bar{e}_n(s)$$

d'après le théorème 13.63.

Désignons par $x = x(s)$ le rayon vecteur de la courbe L et par $\bar{x}(s)$ celui de la courbe $\bar{L}$ (après déplacement). Comme les deux courbes L et $\bar{L}$ ont maintenant pour origine un même point $x(0)$, on a

$$\bar{x}(s) = x(0) + \int_0^s \bar{e}_1(\xi)\, d\xi = x(0) + \int_0^s e_1(\xi)\, d\xi = x(s).$$

Ainsi, la courbe L se confond avec la courbe L.

Les équations

$$\varkappa_1 = \varkappa_1(s), \ \ldots, \ \varkappa_{n-1} = \varkappa_{n-1}(s)$$

s'appellent *équations naturelles de la courbe* L ; nous avons vu qu'elles déterminent la courbe L dans l'espace n-dimensionnel à une transformation linéaire isométrique de cet espace près.

16.42. Courbes à courbures données.

a. Théorème. *Etant données* $n-1$ *fonctions continues quelconques* $\varphi_1(s) > 0, \ \ldots, \ \varphi_{n-1}(s) > 0$ $(0 \leqslant s \leqslant s_0)$, *il est possible de construire une fonction vectorielle* $x = x(s)$ *à valeurs dans* R_n, n *fois continûment dérivable et telle que, pour la courbe* $L \subset R_n$ *définie par l'équation* $x = x(s)$, *les* $\varphi_k(s)$ *représentent les courbures respectives en fonction de la longueur d'arc* s :

$$\varphi_1(s) = \varkappa_1(s), \ \ldots, \ \varphi_{n-1}(s) = \varkappa_{n-1}(s).$$

b. Etablissons d'abord le lemme général suivant :

Lemme. *Considérons dans* R_n *l'équation vectorielle*

$$\frac{dy(t)}{dt} = A(t)\, y(t),$$

où $A(t)$ *est un opérateur antisymétrique* (i.e. tous les éléments de la matrice $A(t) = \| a_{jk}(t) \|$ changent de signe par transposition : $a_{jk}(t) = -a_{kj}(t)$). *La matrice résolvante* $\Omega_{t_0}^t$ *est orthogonale* (i.e. la matrice transposée de $\Omega_{t_0}^t$ est identique à sa matrice inverse).

Démonstration. Nous avons vu dans 13.38e que l'opérateur $\Omega_{t_0}^t$ satisfait à l'équation

$$\frac{d\Omega_{t_0}^t}{dt} = A(t)\, \Omega_{t_0}^t \qquad (1)$$

avec la condition initiale $\Omega_{t_0}^{t_0} = I$. D'autre part, d'après 13.38*d* on a

$$\Omega_{t_0}^{t} \Omega_{t}^{t_0} = I ;$$

en dérivant cette égalité en t on trouve

$$\frac{d\Omega_{t_0}^{t}}{dt} \Omega_{t}^{t_0} + \Omega_{t_0}^{t} \frac{d\Omega_{t}^{t_0}}{dt} = 0$$

ou bien

$$A(t) \Omega_{t_0}^{t} \Omega_{t}^{t_0} + \Omega_{t_0}^{t} \frac{d\Omega_{t}^{t_0}}{dt} = 0,$$

d'où

$$\frac{d\Omega_{t}^{t_0}}{dt} \Omega_{t}^{t_0} = -\Omega_{t}^{t_0} A(t).$$

En passant aux opérateurs adjoints on trouve [14 ; 7.64]

$$\frac{d(\Omega_{t}^{t_0})'}{dt} = -A'(t) (\Omega_{t}^{t_0})'.$$

Utilisons maintenant la condition $A'(t) = -A(t)$; nous avons

$$\frac{d(\Omega_{t}^{t_0})'}{dt} = A(t) (\Omega_{t}^{t_0})'.$$

En comparant cette équation avec (1) et vu que l'opérateur $(\Omega_{t}^{t_0})'$ satisfait à la même condition initiale $(\Omega_{t_0}^{t_0})' = I' = I$ que $\Omega_{t_0}^{t}$ nous obtenons

$$(\Omega_{t}^{t_0})' = \Omega_{t_0}^{t} \tag{2}$$

en vertu de l'unicité de la solution. Or $\Omega_{t}^{t_0} = (\Omega_{t_0}^{t})^{-1}$. L'égalité (2) prouve donc que $\Omega_{t_0}^{t}$ est un opérateur orthogonal, ce qu'il fallait démontrer.

c. Considérons à présent le système d'équations vectorielles

$$\left.\begin{aligned} y_1'(s) &= \varphi_1(s)\, y_2(s), \\ y_2'(s) &= -\varphi_1(s)\, y_1(s) + \varphi_2(s)\, y_3(s), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n'(s) &= -\varphi_{n-1}(s)\, y_{n-1}(s) \end{aligned}\right\} \tag{3}$$

pour les conditions initiales

$$y_1(0) = e_1, \ \ldots, \ y_n(0) = e_n$$

où $e_1, \ldots, e_n$ sont n vecteurs orthogonaux et normés quelconques de l'espace R_n.

Montrons que *les fonctions vectorielles* $y_1(s), \ldots, y_n(s)$ *qui représentent une solution du système* (3) existant dans R_n d'après le théorème 13.63 *sont orthogonales et normées pour tout* $s \in [0, s_0]$.

D'après le lemme *b*, la matrice $\| \omega_{jk}(s) \|$ de l'opérateur résolvant Ω_0^s du système (3) est orthogonale, de sorte que

$$\sum_{j=1}^{n} \omega_{jk}(s)\, \omega_{pk}(s) = \begin{cases} 1 & \text{pour } j = p, \\ 0 & \text{pour } j \neq p. \end{cases}$$

Nous avons

$$y_j(s) = \sum \omega_{jk}(s)\, e_k \qquad (i = 1, \ldots, n).$$

La base $e_1, \ldots, e_n$ étant orthonormée, on a

$$(y_j(s),\ y_p(s)) = \sum_k \omega_{jk}(s)\, \omega_{pk}(s) = \begin{cases} 1 & \text{pour } j = p, \\ 0 & \text{pour } j \neq p, \end{cases}$$

ce qu'il nous fallait.

Ainsi, les vecteurs $y_j(s)$ s'avèrent orthogonaux et normés pour tout $s \in [0, s_0]$.

d. En continuant posons

$$x(s) = \int_0^s y_1(\sigma)\, d\sigma.$$

Nous obtenons une courbe L dans l'espace n-dimensionnel R_n. Comme $| x'(s) | = | y_1(s) | = 1$, le paramètre s est la longueur d'arc de la courbe L. Ensuite, pour tout $m = 1, 2, \ldots, n$, les vecteurs $y_1(s), \ldots, y_m(s)$ sont orthogonaux et normés, $x^{(m)}(s) = y_1^{(m-1)}(s)$ s'exprime linéairement par $y_1(s), \ldots, y_m(s)$ d'après le système (3), le coefficient de $y_m(s)$ étant positif (égal à $\varphi_{m-1}(s)$) ; par conséquent, les vecteurs $y_1(s), \ldots, y_n(s)$ sont pour tout s les vecteurs de la base naturelle de la courbe L. Or, pour les vecteurs d'une base naturelle on a les formules de Frénet 16.27(5)

$$\begin{aligned} y_1'(s) &= \varkappa_1(s)\, y_2(s), \\ y_2'(s) &= -\varkappa_1(s)\, y_1(s) + \varkappa_2(s)\, y_3(s), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_{n-1}'(s) &= -\varkappa_{n-1}(s)\, y_n(s). \end{aligned}$$

En les comparant avec les équations (3) nous établissons successivement les égalités

$$\varphi_1(s) \equiv \varkappa_1(s), \quad \varphi_2(s) \equiv \varkappa_2(s), \ \ldots, \ \varphi_{n-1}(s) \equiv \varkappa_{n-1}(s).$$

Le théorème est démontré.

§ 16.5. Hélices

16.51.a. Définition. On appelle *hélice* une courbe dont toutes les courbures sont constantes.

b. Il est évident qu'une droite satisfait à cette définition (toutes ces courbures sont nulles). Dans un plan, une circonférence Q de rayon R a, nous l'avons vu dans 16.24, la courbure constante $\varkappa = 1/R$; toutes ses courbures supérieures sont nulles. Ainsi, une circonférence satisfait elle aussi à la définition d'une hélice. Montrons qu'il n'existe pas d'autres hélices dans un plan. Si L est une courbe plane d'une courbure constante $\varkappa > 0$, nous considérons à côté d'elle la circonférence Q de rayon $R = 1/\varkappa$ qui a également la courbure constante $\varkappa$. Alors, d'après le théorème 16.41, on peut faire coïncider la courbe L avec la circonférence Q; donc la courbe L est elle-même une circonférence.

c. Dans l'espace tridimensionnel, l'hélice « classique » Q

$$x_1 = a\cos t,\quad x_2 = a\sin t,\quad x_3 = bt, \tag{1}$$

comme nous l'avons vu dans 16.28, a la courbure et la torsion constantes que l'on calcule d'après les formules

$$\varkappa_1 = \frac{a}{a^2+b^2}.\quad \varkappa_2 = \frac{b}{a^2+b^2}. \tag{2}$$

Par conséquent, Q est une hélice au sens de notre définition. Montrons qu'il n'y a pas d'autres hélices dans R_3. Soit L une hélice dans R_3 telle que $\varkappa_1 > 0$, $\varkappa_2 > 0$. En partant de (2) déterminons les paramètres a et b d'après $\varkappa_1$ et $\varkappa_2$ donnés; il est aisé de voir que

$$a = \frac{\varkappa_1}{\varkappa_1^2+\varkappa_2^2}.\quad b = \frac{\varkappa_2}{\varkappa_1^2+\varkappa_2^2}.$$

D'après a et b obtenus, construisons l'hélice (1). Sa courbure est $\varkappa_1$, sa torsion $\varkappa_2$. Vu le théorème 16.41, on peut faire coïncider la courbe L avec l'hélice (1) par un déplacement dans l'espace R_3, ce qu'il nous fallait.

16.52. Trouvons les hélices dans un espace n-dimensionnel. Une hélice dans R_n est une courbe pour laquelle $\varkappa_1(s) = \varkappa_1, \ldots, \varkappa_{n-1}(s) = \varkappa_{n-1}$ soat constantes et non nulles. Les vecteurs $e_1(s), \ldots, e_n(s)$ satisfont au système d'équations de Frénet 16.27(5)

$$\left.\begin{aligned}
e_1'(s) &= \varkappa_1 e_2(s),\\
e_2'(s) &= -\varkappa_1 e_1(s) + \varkappa_2 e_3(s),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
e_{n-1}'(s) &= -\varkappa_{n-2} e_{n-2}(s) + \varkappa_{n-1} e_n(s),\\
e_n'(s) &= -\varkappa_{n-1} e_{n-1}(s)
\end{aligned}\right\} \tag{3}$$

avec la matrice constante des coefficients

$$K = \begin{Vmatrix} 0 & \varkappa_1 & & & & & \\ -\varkappa_1 & 0 & \varkappa_2 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & -\varkappa_{n-2} & 0 & \varkappa_{n-1} \\ & & & & -\varkappa_{n-1} & 0 \end{Vmatrix}$$

Calculons le rang de la matrice K. En barrant la ligne et la colonne contenant l'élément $\varkappa_1$ nous abaissons le rang d'une unité, puis encore d'une unité en barrant la ligne et la colonne contenant l'élément $-\varkappa_1$; nous aboutissons à la matrice

$$\begin{Vmatrix} 0 & \varkappa_3 & & & & \\ -\varkappa_3 & 0 & \varkappa_4 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & -\varkappa_{n-2} & 0 & \varkappa_{n-1} \\ & & & & -\varkappa_{n-1} & 0 \end{Vmatrix}$$

de même structure que la matrice initiale mais de rang inférieur de deux unités. En continuant le procédé deux possibilités se présentent: si $n = 2m$ est pair, le rang vaut n; si $n = 2m + 1$ est impair, nous obtenons en fin de compte la matrice unidimensionnelle d'élément nul, par conséquent le rang de la matrice initiale vaut $2m = n - 1$.

La matrice K est évidemment antisymétrique: elle change de signe par transposition. Utilisons un théorème connu sur la structure d'un opérateur antisymétrique [14; 9.46]. Si $n = 2m$ est pair, il existe dans l'espace R_n une base canonique orthonormée $x_1, y_1, \ldots, x_m, y_m$ telle que

$$Kx_1 = \tau_1 y_1, \quad Kx_2 = \tau_2 y_2, \ \ldots, \quad Kx_m = \tau_m y_m,$$
$$Ky_1 = -\tau_1 x_1, \quad Ky_2 = -\tau_2 x_2, \ \ldots, \quad Ky_m = -\tau_m x_m.$$

Pour $n = 2m + 1$, il existe encore un vecteur de base z_n pour lequel

$$Kz_n = 0.$$

Comme le rang de la matrice K vaut $2m$, tous les nombres $\tau_1, \ldots, \tau_m$ sont dans le présent cas non nuls.

Parallèlement au système vectoriel (3), considérons le système scalaire

$$\left.\begin{aligned} u_1'(s) &= \varkappa_1 u_2(s), \\ u_2'(s) &= -\varkappa_1 u_1(s) + \varkappa_2 u_3(s), \\ &\cdots\cdots\cdots\cdots \\ u_n'(s) &= -\varkappa_{n-1} u_{n-1}(s). \end{aligned}\right\} \tag{4}$$

Il possède n solutions vectorielles linéairement indépendantes qui sont respectivement colinéaires, pour $s = 0$, aux vecteurs canoniques de la matrice K.

Construisons la matrice de l'opérateur résolvant e^{tK}. En vertu de 13.14c (et 13.14a), nous avons dans la base $x_1, y_1, x_2, y_2, \ldots, x_m, y_m$ (z_n):

$$e^{tK} = \begin{Vmatrix} \cos t\tau_1 & -\sin t\tau_1 & \ldots & & \\ \sin t\tau_1 & \cos t\tau_1 & \ldots & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & \ldots & \cos t\tau_m & -\sin t\tau_m \\ & & \ldots & \sin t\tau_m & \cos t\tau_m \end{Vmatrix}.$$

Si l'on pose

$$x_j=(x_{j1}, \ldots, x_{jn}),\ y_j=(y_{j1}, \ldots, y_{jn}),$$
$$x_j(s)=(x_{j1}(s), \ldots, x_{jn}(s)),\ y_j(s)=(y_{j1}(s), \ldots, y_{jn}(s)), \qquad (5)$$
$$j=1, \ldots, n,$$
$$z_n=(z_{n1}, \ldots, z_{nn}),$$
$$z_n(s)=(z_{n1}(s), \ldots, z_{nn}(s)).$$

alors les formules 13.15*c* (et 13.15*a*) donnent

$$x_{jk}(s)=\cos\tau_j s\cdot x_{jk}+\sin\tau_j s\cdot y_{jk}, \quad (k=1, \ldots, n,\ j=1, \ldots, m)$$
$$y_{jk}(s)=-\sin\tau_j s\cdot x_{jk}+\cos\tau_j s\cdot y_{jk},$$
$$z_{nk}(s)=z_{nk} \qquad (n=2m+1).$$

Soit maintenant $e_1(s), \ldots, e_n(s)$ une solution du système (3) que l'on peut noter

$$e_j(s)=(e_{j1}(s), \ldots, e_{jn}(s)).$$

Pour p fixe, les fonctions $e_{jp}(s)$ ($j=1, \ldots, n$) vérifient le système (4). Soumettons la solution $(e_{jp}(s))$ aux conditions initiales

$$e_{11}(0)=x_{11}, \ldots, e_{n1}(0)=x_{1n},$$
$$e_{12}(0)=y_{11}, \ldots, e_{n2}(0)=y_{1n},$$
$$e_{13}(0)=x_{21}, \ldots, e_{n3}(0)=x_{2n},$$
$$e_{14}(0)=y_{21}, \ldots, e_{n4}(0)=y_{2n},$$
$$\ldots\ldots\ldots\ldots$$
$$e_{1n}(0)=z_{n1}, \ldots, e_{nn}(0)=z_{nn} \quad (n=2m+1).$$

Les vecteurs $e_1(0), \ldots, e_n(0)$ sont normés et orthogonaux (ce qui est nécessaire pour restituer une courbe d'après les courbures données, 16.42*c*). Nous avons en particulier :

$$e_1(s)=(e_{11}(s), \ldots, e_{1n}(s))=(x_{11}(s), y_{11}(s), x_{21}(s), y_{21}(s), \ldots, (z_{n1}(s)))=$$
$$=(\cos\tau_1 s\cdot x_{11}+\sin\tau_1 s\cdot y_{11}, \ -\sin\tau_1 s\cdot x_{11}+\cos\tau_1 s\cdot y_{11}, \ \ldots, \ (z_{n1})).$$

La quantité z_{n1} n'est considérée que pour $n=2m+1$.

En intégrant par rapport à s on obtient (les constantes d'intégration qui représentent les translations le long de chaque axe sont choisies nulles) :

$$r(s)=\left(\frac{\sin\tau_1 s}{\tau_1}x_{11}-\frac{\cos\tau_1 s}{\tau_1}y_{11}, \ \frac{\cos\tau_1 s}{\tau_1}x_{11}+\frac{\sin\tau_1 s}{\tau_1}y_{11}, \ \ldots, (z_{n1}s)\right).$$

On peut écrire ces formules plus brièvement :

$$r(s)=(A_1\cos\tau_1(s-s_1), \ A_1\sin\tau_1(s-s_1),$$
$$A_2\cos\tau_2(s-s_2), \ A_2\sin\tau_2(s-s_2), \ \ldots, (C_n s)). \qquad (6)$$

Pour n pair, $n=2m$, toute la courbe L se trouve évidemment sur la sphère

$$x_1^2+x_3^2+x_5^2+\ldots+x_{2m-1}^2+x_{2m}^2=A_1^2+A_2^2+\ldots+A_m^2;$$

la courbe est fermée si tous les nombres $\tau_1, \ldots, \tau_m$ sont commensurables (i.e. sont des multiples rationnels de l'un d'eux) et ne l'est pas (et n'a pas de point double) si au moins un couple de nombres τ_j, τ_k n'est pas commensurable. Pour $n=2m+1$, la courbe (6) n'est pas bornée ; elle va vers l'infini en coordonnée x_{2m+1} lorsque $t\to+\infty$.

16.53. Hélices dans un espace de dimension infinie. Notons d'abord le fait suivant. Soit L une courbe dans R_n. Supposons que, quels que soient un point donné A (origine) et un autre point P sur cette courbe, il existe un déplacement de l'espace (suivi peut-être d'une symétrie) qui transforme la courbe L en elle-même en conservant le sens de variation du paramètre et en faisant passer le point A en P. Alors, un déplacement de l'espace conservant la métrique, toutes les courbures de la courbe L aux points A et P sont respectivement égales (si la courbe est suffisamment lisse). Comme le point P est arbitraire, toutes les courbures de la courbe L sont constantes; par conséquent, nous avons affaire à une hélice. Inversement, quels que soient deux points A et P d'une hélice L donnée dans R_n, il existe en vertu de 16.41 un déplacement de l'espace R_n (suivi peut-être d'une symétrie) qui transforme l'hélice L en elle-même en conservant le sens de variation du paramètre et en transformant le point A en P.

Ainsi, nous sommes amenés à une nouvelle définition d'une hélice : c'est une courbe L que l'on peut transformer en elle-même par un déplacement de l'espace (avec, peut-être, une symétrie) qui conserve le sens de variation du paramètre et transforme un point A donné sur L en un autre point P donné sur L. Remarquons que cette définition n'exige de L aucune dérivabilité. Une courbe L possédant la propriété mentionnée sera dite *autocongruente*. On peut démontrer que, dans R_n, la classe des hélices se confond avec celle des courbes autocongruentes (cf. exercice 2).

Il s'avère que, dans un espace de dimension infinie, il existe des courbes autocongruentes continues, mais sans tangente. Nous nous bornons ici à un seul exemple de courbe autocongruente dans l'espace hilbertien, qui est tout de même bien caractéristique.

Exemple (« spirale de Wiener »). Considérons l'espace $H_2(0, \infty)$ formé des fonctions $x(\tau)$ réelles continues par morceaux sur la demi-droite $0 \leqslant \tau < \infty$ dont chacune est nulle à l'extérieur d'un intervalle $[0, a]$ (dépendant de la fonction $x(\tau)$). Munissons cet espace d'un produit scalaire et d'une norme d'après les formules

$$(x(\tau),\ y(\tau)) = \int_0^{\infty} x(\tau)\, y(\tau)\, d\tau,$$

$$\| x(\tau) \|^2 = \int_0^{\infty} x^2(\tau)\, d\tau ;$$

les intégrales ont ∞ pour limite supérieure, mais en fait elles sont prises sur un intervalle fini. Pour tout $t \in [0, \infty)$, considérons un élément de l'espace $H_2(0, \infty)$ suivant la règle :

$$Z(t) \equiv s(t, \tau) = \begin{cases} 1 \text{ pour } 0 \leqslant \tau \leqslant t, \\ 0 \text{ pour } \tau > t. \end{cases} \tag{1}$$

Lorsque t varie de 0 à ∞, le point $Z(t)$ décrit dans l'espace $H_2(0, \infty)$ une courbe L qui s'appelle spirale de Wiener. Cette courbe est continue (même uniformément) puisque

$$\| Z(\bar{t}) - Z(t) \|^2 = \left| \int_t^{\bar{t}} 1^2 \cdot d\tau \right| = |\bar{t} - t|.$$

Montrons que la courbe L est autocongruente. Soit un t_0, $0 < t_0 < \infty$. Considérons la transformation U de l'espace $H_2(0, \infty)$ dans lui-même définie par

la formule

$$x(\tau) \to Ux(\tau) = z(t_0, \tau) + x(\tau - t_0)$$

(pour $\tau < t_0$ nous posons $x(\tau - t_0) = 0$).

Cette transformation est isométrique car

$$\| Ux(\tau) - Uy(\tau) \|^2 = \| x(\tau - t_0) - y(\tau - t_0) \|^2 =$$

$$= \int_{t_0}^{\infty} [x(\tau - t_0) - y(\tau - t_0)]^2 \, d\tau = \int_{0}^{\infty} [x(\tau) - y(\tau)]^2 \, d\tau = \| x(\tau) - y(\tau) \|^2.$$

Le point origine $Z(0)$ de la courbe se confond avec le zéro de l'espace $H_2(0, \infty)$. La transformation U le transforme en le point $Z(t_0)$. Tout point $Z(t)$ de la courbe L est transformé par U en le point $Z(t + t_0)$ de la même courbe. La courbe L est donc autocongruente.

Montrons à présent que la fonction $Z(t)$ n'a pas de dérivée dans l'espace $H_2(0, \infty)$. En effet, le vecteur

$$\frac{Z(t+h) - Z(t)}{h} = \frac{z(\tau, t+h) - z(\tau, t)}{h} \tag{2}$$

correspond à la fonction de τ qui vaut 0 pour $\tau \notin [t, t+h]$ et $1/h$ pour $\tau \in [t, t+h]$. Sa norme est $1/\sqrt{h}$, de sorte que le rapport (2) n'a aucune limite lorsque $h \to 0$.

La courbe L possède encore une propriété intéressante : *ses deux cordes quelconques correspondant à deux intervalles disjoints de variation du paramètre sont orthogonales entre elles*. En effet,

$$(Z(t+h) - Z(t),\ Z(z+k) - Z(s)) =$$

$$= \int_{0}^{\infty} [z(t+h, \tau) - z(t, \tau)] [z(s+k, \tau) - z(s, \tau)] \, d\tau = 0,$$

si les intervalles $(t, t+h)$ et $(s, s+k)$ sont disjoints.

On peut construire des analogues lisses de la spirale de Wiener si l'on remplace la fonction $z(\tau, t)$ de (1) par la translatée d'un nombre t le long de l'axe des τ d'une fonction dérivable fixe $\varphi_0(\tau)$.

Pour de telles courbes lisses, on peut calculer les courbures selon nos règles habituelles ; bien entendu, toutes les courbures sont constantes.

Exercices

1. Démontrer que le lieu géométrique des centres de courbure d'une hélice dans R_3 est encore une hélice de même axe ; le lieu géométrique de ses centres de courbure est l'hélice initiale.

2. Démontrer que toute courbe autocongruente dans R_n est une hélice (sans supposer la dérivabilité continue).

3. On établit une correspondance biunivoque entre les points de deux courbes dans R_n d'une telle manière qu'aux points correspondants les vecteurs des bases naturelles sont respectivement parallèles. Soient $\varkappa_j^{(1)}$, $\varkappa_j^{(2)}$ ($j = 1, 2, \ldots, n-1$) les courbures de ces courbes ; montrer que

$$\frac{\varkappa_1^{(1)}}{\varkappa_1^{(2)}} = \frac{\varkappa_2^{(1)}}{\varkappa_2^{(2)}} = \ldots = \frac{\varkappa_{n-1}^{(1)}}{\varkappa_{n-1}^{(2)}}.$$

4. La sphère osculatrice m-dimensionnelle S_m à une courbe gauche L est la sphère dans l'espace $(m+1)$-dimensionnel défini par les $m+1$ premiers vecteurs de la base naturelle et telle que l'écart d'un point de la courbe L à cette sphère a Δs^{m+2} pour ordre de petitesse. Montrer que le rayon r_m de la sphère osculatrice m-dimensionnelle ne diminue pas avec la croissance de m.

5. (« Courbe de Peano »). Soit $t = 0, t_1 t_2 \ldots t_{2n-1} t_{2n} \ldots$ le développement ternaire formé des symboles 0 et 2 d'un nombre $t \in [0, 1)$. Montrer que les fonctions $x(t) = 0, t_1 t_3 \ldots t_{2n-1} \ldots$ et $y(t) = 0, t_2 t_4 \ldots t_{2n} \ldots$ sont définies d'une façon univoque sur l'ensemble de tous les t de forme indiquée et qu'elles admettent un prolongement continu sur l'intervalle $[0, 1]$. Montrer que la courbe $r(t) = \{x(t), y(t)\}$ passe par tous les points du carré $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$.

Historique

Pour l'espace tridimensionnel, les équations fondamentales de la théorie des courbes sont données par Serret (1851) et Frénet (1852); Jordan généralise ces équations au cas n-dimensionnel (1874). Les hélices dans R_n sont décrites par Forsythe (1930). Les hélices autocongruentes dans l'espace hilbertien jouent un rôle important dans la théorie moderne des probabilités où elles sont appelées « processus stochastiques d'accroissements stationnaires » (et décrivent certains phénomènes réels, par exemple, le mouvement brownien, le mouvement turbulent de fluides, etc.; cf. à ce propos [17] et [7]). Plusieurs grands savants de notre époque, Wiener, von Neumann, Kolmogorov, M. Krein, étudient diverses propriétés de ces courbes. Kolmogorov indique la forme canonique d'une courbe autocongruente dans l'espace hilbertien, Krein découvre qu'un « arc hélicoïdal », i.e. une partie finie d'une courbe autocongruente, peut être prolongé de différentes façons jusqu'à former une courbe autocongruente complète et classe tous les prolongements possibles. Les travaux des auteurs cités sur les courbes autocongruentes dans l'espace hilbertien se rapportent aux années 1939-1943.

Indications et réponses

Chapitre §2

§. *Réponse.* Oui dans les cas a) et b), non dans le cas c).

2. *Indication.* L'ensemble de toutes les suites croissantes de nombres naturels a la puissance du continu (chap. 2, exercice 8). Pour chaque telle suite $n_1 < n_2 < \ldots$, choisir la fonction $x(t) \in R^a(0, \infty)$ qui vaut l'unité aux points $n_1, n_2, \ldots$ et est nulle aux autres points entiers.

3. *Indication.* Si une fonction extrémale existait, elle prendrait la valeur 1 pour $0 < x < 1/2$ et -1 pour $1/2 < x < 1$.

4. *Indication.* Il faut prouver que (x, y) vérifie les axiomes du produit scalaire 12.41. Pour l'axiome 12.41*d*, appliquer le lemme sur le parallélogramme aux parallélogrammes construits sur les vecteurs $x + z, y$; $x - z, y$; $y + z, x$; $y - z, x$. Pour l'axiome 12.41*c*, considérer d'abord α entiers, puis fractionnaires, enfin quelconques en passant à la limite.

5. *Indication.* D'après le théorème de Cauchy, on a $\int_{|z|=1} p(z)\,dz = 0$. Cette égalité se conserve par le passage à la limite.

6. *Indication.* Poser $F = \bigcap_{f \in I} \{x \in Q : f(x) = 0\}$. Montrer que toute fonction $g(x) \in R^a(Q)$ qui est nulle au voisinage de l'ensemble F appartient à l'idéal I. Ensuite, toute fonction $f(x) \in R^a(Q)$ qui est nulle sur F est la limite de fonctions de la forme $g(x)$.

7. *Indication.* Soit $\delta > 0$ un nombre qui correspond au nombre $\varepsilon > 0$ d'après la condition d'équicontinuité de la famille E; alors les valeurs des fonctions $x(t) \in E$ aux points d'un δ-réseau fini du compact Q forment un 2ε-réseau précompact pour l'ensemble E.

8. *Indication.* Sans restreindre la généralité, on peut supposer que $\left|\sum_{m=1}^{\infty} t_{km} - 1\right| < \delta$ quel que soit $k = 1, 2, \ldots$. Définir une suite numérique $N_1, N_2, \ldots$ et une suite numérique $k_1, k_2, \ldots$ de façon à avoir

$$\sum_{m=N_1}^{\infty} |t_{1m}| < \delta;\quad \sum_{m=1}^{N_1} |t_{k_1 m}| < \delta;\quad \sum_{N_2}^{\infty} |t_{k_1 m}| < \delta;\quad \sum_{1}^{N_2} |t_{k_2 m}| < \delta, \ldots,$$

poser $\xi_n = 1$ pour $N_{2p-1} \leqslant n < N_{2p}$ et -1 pour $N_{2p} \leqslant n < N_{2p+1}$, $p = 1, 2, \ldots$

9. *Indication.* Il suffit de considérer les suites $x = \{\xi_n\}$ pour lesquelles $\sup \xi_n = \overline{\lim}\, \xi_n = -\underline{\lim}\, \xi_n = -\inf \xi_n$.

10. *Indication.* Utiliser le théorème d'unicité et le principe du maximum pour les fonctions analytiques.

11. *Indication.* L'opérateur de multiplication par $x - \lambda$ ne peut avoir pour inverse que l'opérateur de multiplication par $1/(x - \lambda)$.

12. *Indication.* On peut tirer de 12.67 que le spectre de l'opérateur A est composé par les seules valeurs propres généralisées. Ensuite, on doit déduire $(p(A) - p(\lambda) E) x_n \to 0$ de la relation $(A - \lambda E) x_n \to 0$, $\| x_n \| = 1$ et utiliser la compacité de $p(A)$ et la condition $p'(\lambda) \neq 0$.

13. *Indication.* Utiliser l'exercice 12.

14. *Indication.* Il suffit de considérer le cas

$$\int_a^b |x(t)|^p \, dt = \int_a^b |y(t)|^q \, dt = 1.$$

Intégrer l'inégalité de Young (9.61*h*)

$$|x(t)\, y(t)| \leqslant \frac{1}{p} |x(t)|^p + \frac{1}{q} |y(t)|^q.$$

15. *Indication.* Intégrer l'inégalité

$$|f(x) + g(x)|^p \leqslant (|f(x)| + |g(x)|)^p \leqslant$$
$$\leqslant |f(x)| (|f(x)| + |g(x)|)^{p-1} + |g(x)| (|f(x)| + |g(x)|)^{p-1}$$

et appliquer l'inégalité de Hölder dans le deuxième membre (exercice 14).

16. *Indication.* Utiliser la méthode de l'exercice 14 en remplaçant l'intégration par la sommation.

17. *Indication.* Analogie avec l'exercice 15.

18. *Indication.* Si $x_n = (\xi_{n1}, \ldots, \xi_{nk}, \ldots) \in l_p$ est de Cauchy, alors, pour tout $k = 1, 2, \ldots$, la suite numérique ξ_{nk} $(n = 1, 2, \ldots)$ est également de Cauchy; soit $\xi_k = \lim_{n \to \infty} \xi_{nk}$. Pour tout $\varepsilon > 0$, il existe un N tel que, quels que soient n, $m \geqslant N$, on a l'inégalité $\sum_{k=1}^{\infty} |\xi_{nk} - \xi_{mk}| \leqslant \varepsilon$. Remplacer ici ∞ par p, passer à la limite pour $m \to \infty$, puis pour $p \to \infty$.

19. *Indication.* En appliquant le théorème d'Arzelà (12.24*d*) obtenir une sous-suite uniformément convergente; en appliquant le théorème d'Arzelà encore une fois obtenir une suite plus raréfiée avec les dérivées uniformément convergentes, etc.; ensuite considérer la sous-suite diagonale.

20. *Indication.* L'ensemble $\{x \in R_n : \| x \|_p \leqslant 1\}$ n'est pas convexe.

21. *Indication.* Pour l'ensemble $E \subset P(Q)$ formé d'une seule fonction $x(t)$, on peut poser $Q_k = \{t \in Q : k\varepsilon < x(t) \leqslant (k+1)\varepsilon\}$. Dans le cas général, appliquer le critère de Haudsdorff 3.93*c*.

Chapitre 13

1. *Indication.* Vérifier la condition de Lipschitz.

2. *Réponse.* (a) parabole; (b) parabole semi-cubique.

3. *Indication.* Utiliser l'expression du wronskien (13.74).

4. *Indication.* Dans une base jordanienne, le système se décompose en autant de systèmes indépendants qu'il y a de racines caractéristiques. Chaque système est équivalent à une équation de la forme $\left(\frac{d}{dt} - \lambda\right)^m u(t) = 0$. Le système en entier est équivalent à l'équation

$$\prod_k \left(\frac{d}{dt} - \lambda_k\right)^{m_k} u(t) = 0.$$

5. *Indication.* Les opérateurs Ω_0^t et Ω_T^{t+T} vérifient une même équation et une même condition initiale.

6. *Indication.* Définir l'opérateur A (t) sur le sous-espace X_t engendré par les vecteurs $u_1(t), \ldots, u_n(t)$ par les conditions $u_k'(t) = A(t)\, u_k(t)$ ($k = 1, \ldots, n$) et sur le supplémentaire orthogonal de X_t poser A (t) nul. Vérifier la continuité de A (t) en t.

7. *Réponse.* C'est possible. En tant qu'exemple pour $n = 2$, on peut choisir les fonctions

$$y_1(t) = \begin{cases} t^3 \text{ pour } t > 0 \\ 0 \text{ pour } t < 0 \end{cases}, \quad y_2(t) = y_1(-t).$$

8. *Réponse.* Par exemple,

$$\begin{vmatrix} y^{(n)}(t) & y^{(n-1)}(t) & \ldots & y(t) \\ y_1^{(n)}(t) & y_1^{(n-1)}(t) & \ldots & y_1(t) \\ \ldots & \ldots & \ldots & \ldots \\ y_n^{(n)}(t) & y_n^{(n-1)}(t) & \ldots & y_n(t) \end{vmatrix} = 0.$$

9. *Indication.* Pour $m = 0$ le résultat découle du théorème d'existence. Ensuite on procède par récurrence.

10. *Indication.* Substitution $y(t) = e^{kt} z(t)$.

11. *Indication.* Substitution $y(t) = k \int_0^t w(s)\, ds$.

12. *Indication.* Une ε-presque-solution vérifie également l'inégalité

$$\left\| y(t) - y(0) - \int_0^t f(s, y(s))\, ds \right\| \leqslant \varepsilon t.$$

Utiliser la solution de l'exercice 11.

13. *Indication.* Pour $t_k \leqslant t \leqslant t_{k+1}$ on a

$$y_\Pi(t) = \left\{ (E + (t - t_k) A(t_k)) \prod_{j=k-1}^{0} (E + A(t_j)\, \Delta t_j) \right\} y_0,$$

$$y_\Pi'(t) = \left\{ A(t_k) \prod_{j=k-1}^{0} (E + A(t_j)\, \Delta t_j) \right\} y_0 =$$

$$= A(t_k) [E + (t - t_k) A(t_k)]^{-1} y_\Pi(t) =$$

$$= A(t_k) [E + B_k(t)]\, y_\Pi(t) A(t)\, y_\Pi(t) + C_k(t)\, y_\Pi(t),$$

où les opérateurs $B_k(t)$ et $C_k(t)$ tendent vers zéro pour un morcellement illimité de la partition II.

14. *Indication.* Utiliser les solutions des exercices 13 et 12.

15. *Indication.* Appliquer la méthode de l'exercice 13.

16. *Indication.* Utiliser les solutions des exercices 15 et 12.

Chapitre 14

1. *Indication.* Substituer dans les séries trouvées certaines valeurs numériques de la variable.

2. *Indication.* Ibid.

3. *Réponse.* (a) $s(x) = e^{\cos x} \cos(x \sin x)$,

(b) $s(x) = e^{\cos x} \sin(x \sin x)$.

4. *Indication.* Cet ensemble précompact possède un seul point limite.

5. *Indication.* Appliquer le second théorème de la moyenne (exercice 3 du chap. 9).

6. *Indication.* Préciser convenablement la solution de l'exercice 5.

7. *Indication.* Pour n impair on a

$$\frac{4}{\pi} b_n = \int_0^{\pi/2} f(t) \sin nt\, dt = \int_0^{\pi/n} f(t) \sin nt\, dt + \int_{\pi/n}^{\pi/2} f(t) \sin nt\, dt,$$

et il suffit de démontrer que

$$I_1 = \int_0^{\pi/n} f(t) \sin nt\, dt = \frac{\beta_n}{n} f\left(\frac{\pi}{n}\right), \quad \beta_n \to 2,$$

$$I_2 = \int_{\pi/n}^{\pi/2} f(t) \sin nt\, dt = -\frac{1}{n} f\left(\frac{\pi}{n}\right) + \gamma_n, \quad \sum_1^{\infty} |\gamma_n| < \infty.$$

Dans le premier cas on remplace nt par τ, dans le deuxième on intègre par parties avec l'application du théorème de la moyenne et de l'exercice 16 du chap. 7.

8. *Indication.* Ajouter $\sum_1^{\infty} \frac{1}{n} f\left(\frac{\pi}{n}\right) = \infty$ aux conditions 1)-4) de l'exercice 7. Utiliser l'exercice 6.

9. *Indication.* Mettre sous forme complexe la série de Fourier pour la fonction $f(t)$ de l'exercice 8.

10. *Indication.* $(xQ_n, Q_k) = (Q_n, xQ_k)$ et xQ_k, pour $k < n - 1$, est un polynôme de degré inférieur à n ; par conséquent, on a

$$xQ_n(x) = \gamma_{n+1} Q_{n+1}(x) + \gamma_n Q_n(x) + \gamma_{n-1} Q_{n-1}(x) \qquad (1)$$

avec des constantes γ_{n+1}, γ_n, γ_{n-1}. Comparer les coefficients de x^{n+1}, x^n et calculer (xQ_n, Q_{n-1}).

11. *Indication.* Multiplier (1) par $Q_n(t)$. En remplaçant t par x et x par t soustraire l'identité obtenue de la précédente. Sommer en n.

12. *Indication.* La première propriété découle de l'orthogonalité à (1), la deuxième de l'orthogonalité au polynôme $\prod_{k=1}^{m} (x - x_k)$ dans la supposition que $x_1, \ldots, x_m$ soient *toutes* les racines du polynôme $Q_n(t)$ et que $m < n$.

Chapitre 15

1. *Indication.* Appliquer la méthode des exercices 5 et 6 du chap. 14.

2. *Indication.* Utiliser l'idée de la résolution de l'exercice 8 du chap. 14.

3. *Indication.* Utiliser l'idée de la résolution de l'exercice 9 du chap. 14.

4. *Indication.* Vérifier l'affirmation pour les fonctions de la classe S (15.29*a*) ; puis pour la fonction $f_N(x)$ qui vaut $f(x)$ pour $|x| \leqslant N$ et zéro pour $|x| > N$ en l'approchant par les fonctions de la classe S ; effectuer le passage à la limite lorsque $N \to \infty$.

5. *Indication.* Considérer l'intégrale $\int_{-\infty}^{\infty} |tx\varphi(x) + \varphi'(x)|^2 dx \geqslant 0$ comme trinôme du second degré par rapport au paramètre t.

6. *Réponse.* $F_1(p) = F(p-a)$, $F_2(p) = pF(p)$,

$$F_3(p) = \frac{1}{p} F(p), \quad F_4(p) = -F'(p), \quad F_5(p) = \int_p^\infty F(s)\, ds,$$

où l'intégrale est prise le long de n'importe quel chemin s'éloignant à l'infini dans le demi-plan $\operatorname{Re} p > \gamma_0$.

7. *Réponse.*

$$\Phi_1(p) = \frac{1}{p-a}, \quad \Phi_2(p) = \frac{\Gamma(\alpha)}{p^\alpha},$$

$$\Phi_3(p) = \frac{\Gamma(\alpha)}{(p-a)^\alpha}, \quad \Phi_4(p) = \frac{a}{p^2+a^2},$$

$$\Phi_5(p) = \frac{p}{p^2+a^2}, \quad \Phi_6(p) = \frac{1}{a} \operatorname{arctg} \frac{a}{p}$$

(pour $p > 0$).

8. *Indication.* Poser $z = e^t$.

Chapitre 16

1. *Indication.* Pour l'hélice $r = (a\cos t,\ a \sin t,\ bt)$, l'hélice des centres de courbure a b^2/a pour rayon de sa projection sur le plan horizontal.

2. *Indication.* Les transformations orthogonales d'autocongruence commutent, donc possèdent une base canonique commune.

3. *Indication.* Tous les rapports indiqués valent $ds^{(1)}/ds^{(2)}$.

4. *Indication.* $S_{m-1} \subset S_m$.

5. *Indication.* Utiliser les développements décimaux des deux coordonnées d'un point donné du carré.

Bibliographie

[1] М. С. Бродский, Треугольные и жордановы представления линейных операторов, «Наука», 1969.

[2] А. Л. Брудно, Суммирование ограниченных последовательностей матрицами, Матем. сб., т. 16, стр. 191-245, 1945.

[3] И. М. Гельфанд, Д. А. Райков и Г. Е. Шилов, Коммутативные нормированные кольца, Физматгиз, 1960.

[4] И. Ц. Гохберг и М. Г. Крейн, Введение в теорию линейных несамосопряженных операторов, «Наука», 1965.

[5] И. Ц. Гохберг и М. Г. Крейн, Теория вольтерровых операторов в гильбертовом пространстве и ее приложения, «Наука», 1967.

[6] С. Мандельбройт, Квазианалитические классы функций, Гостехиздат, 1936.

[7] А. С. Монин и А. М. Яглом, Статистическая гидромеханика, «Наука», 1967, т. 2, гл. 8.

[8] М. А. Наймарк, Нормированные кольца, изд. 2-е, «Наука», 1968.

[9] И. П. Натансон, Конструктивная теория функций, М.-Л., 1949,

[10] В. П. Паламодов, Линейные дифференциальные операторы с постоянными коэффициентами, «Наука», 1967.

[11] И. Г. Петровский, Лекции об уравнениях с частными производными, Физматгиз, 1961 (3. изд.), гл. II, § 9; гл. III, § 28; гл. IV.

[12] Г. М. Фихтенгольц, Курс дифференциального и интегрального исчисления, т. 3, Гостехиздат, 1949, гл. XIX, § 5.

[13] Г. Е. Шилов, Математический анализ, Второй специальный курс, «Наука», 1965.

[14] Г. Е. Шилов, Математический анализ. Конечномерные линейные пространства, «Наука», 1969.

[15] Г. Е. Шилов, Математический анализ, Специальный курс, Физматгиз, 1961; стр. 262.

[16] Г. Е. Шилов и Б. Л. Гуревич, Интеграл, мера и производная, М., 1967 гл. 2.

[17] А. М. Яглом, Случайные процессы со стационарными приращениями, Успехи мат. наук, 1952, т. 7, № 5.

[18] R. Agnew, Equivalence of methods for evaluation sequences, Proc. Amer. Math. Soc., t. 3, S. 550-565, 1952.

[19] G. Alexits, Convergence problems of orthogonal series. Lnd (a.o.), Pergamon Press, 1961.

[20] S. Banach, Théorie des opérations linéaires, N.Y. Chelsea publ. co., 1955.
[21] R. Cooke, Infinite matrice and sequence spaces. Lnd., Macmillan 1950.
[22] R. Courant and D. Hilbert, Methods of mathematical physics, vol. 1, New York-London 1955.
[23] I. Guelfand et G. Chilov, Les distributions, Dunod, 1962-1967.
[24] D. Jackson, Fourier series and orthogonal polynomials, 3rd. ed., Carus Mathematical Monographs, Oberlin (Ohio) 1946.
[25] St. Kaczmarz und H. Steinhaus, Theorie des Orthogonalreihen, Warszawa-Lwow 1935.
[26] V. Smirnov, Cours de mathématiques supérieures, t. III, 2e partie, VI-1-6, Editions de Moscou 1972.

Index

Algèbre 25
  commutative 25
  de Gelfand 115
  normée 9, 114
  quotient 26
Alternative de Fredholm 127
Application 147
  contractante 146
Arzelà 132
Ascoli 132

Banach 133, 171
Base 13
  jordanienne 24
  — réelle 24
  naturelle 271
  orthonormée 62
Bernoulli D. 216
Bourbaki 7
Bromwich 260

Carleman 260
Carson 260
Case jordanienne 23
Cauchy 133, 171, 218, 260
Centre de courbure 271
Cercle osculateur 271
Cesàro 133
Classe
  d'équivalence 15
  S 236
  W_M 237
  $W\Omega$ 237
Classes quasi analytiques 250
Coefficients de Fourier 175
Complété
  d'un espace hilbertien 65
  — normé 53
Condition de Dini 166
  unilatérale 189
Condition de Lipschitz 150
  d'ordre α 186
Convolution de fonctions 238
Corde 200
Corps 27
  symétrique 29
Courbe
  autocongruente 285
  de Pesno 267
  dans R_n 261
Courbure 270
Courbures d'ordre supérieur 273
Critère
  d'Abel-Dirichlet 51
  de Cauchy 51
  de Wolerstraes 51

d'Alembert 218
Demi-droite de Valiron 251
Denjoy 260
Dépendance linéaire 12
Déterminant de Wronski 164
Développement d'un vecteur suivant une base 13
Dini 132, 218
Dirac 76
Dirichlet 216
Distance 31
Diviseur généralisé de zéro 115
Doetsch 260
Du Bois-Reymond 218

Egalité de Parseval 175
Elément
  inverse 25
  opposé 10
Engels 261
Ensemble
  absolument convexe 55
  convexe 45
  équilibré 45
  partout dense 33
Enveloppe convexe 86
  fermée 87
Epimorphisme 16
  de l'algèbre 26
Equation(s)
  caractéristique 144
  différentielle 134
  — linéaire homogène 137
  — — non homogène 166
  naturelles 279

Espace
de Banach 41
complexe 10
dual 103
de Hilbert 9, 56
— complexe 63
métrique 9, 31
— compact 35
— précompact 35
— séparable 33
normé 9, 41
des opérateurs linéaires 22, 102
préhilbertien 66
quotient 15
réel 10
vectoriel 9, 10
— affine 9
— complexe 10
— de dimension infinie 13
— — n 13
— normé 9, 41
— — complexe 54
— — réel 41
— réel 10
$C^n(M)$ 12
$D_m(a, b)$ 36
K_n 11
$K^n(E)$ 12
K (E) 12
l_p 44
$L^p(a, b)$ 39
$L^p_p(a, b)$ 39
P^p (M) 33
$R(E)$ 11
R (E) 11
$R^n(M)$ 12
$R^n_n(M)$ 12
Euler 139, 171, 218, 260

Famille séparatrice des points 68
Fonction
duale 229
entière de type exponentiel fini 231
séparatrice des points 68
à valeurs dans un espace normé 61-63, 84, 86, 87, 90, 93
Fonctionnelle linéaire 17
Formule d'inversion
de Fourier 220
de Laplace 243
de Mellin 260
Formule de Taylor 91
Formules de Frénet 273

Forsythe 287
Fourier 172, 216, 260
Fredholm 132
Frénet 287

Gelfand 139
Grassmann 132

Hahn 133
Heavyside 219, 260
Hélice
dans un espace de dimension infinie 285
dans R_3 275
dans R_n 282
Hilbert 9, 132, 133, 218
Hobson 218

Idéal 26
Identité de Christoffel-Darboux 214
Inégalité
de Bessel 174, 175
de Hölder 131
de Young 229
Intégrale
de Fourier 220
—, formule d'inversion 220
multiplicative 146
de Poisson 241
Isomorphisme 16
de l'algèbre 26

Jordan 287

Keldych 137
Kolmogorov 287
Krein 287

Lagrange 171
Laplace 134, 216, 260
Legendre 216
Leibniz 171
Lemme sur le parallélogramme 58
Limite généralisée 110
de Cesàro 110
de Toeplitz 112
de Voronoï 112
Lipschitz 171
Livchitz 133
Longueur d'arc 261

Matrice jordanienne 24
réelle 24
Matrice de Wronski 164
Membrane 202
Monomorphisme 16
Morphisme 15
de l'algèbre 26
Multiplication des opérateurs 17

Neumann 218
Neumann von 133, 287
Newton 171
Norme
d'un opérateur linéaire 95
d'un vecteur 41
Normes équivalentes 46
Noyau
de Dirichlet 189
de Fejér 208
— pour l'intégrale de Fourier 225
de Fourier-Legendre 214
de Poisson 204

Opérateur
compact 124
de Fredholm 20, 127
inverse 18
linéaire 17
— borné 94
— compact 124
— continu 94
résolvant 152
— d'une équation linéaire 163
de Volterra 129
Orthogonalisation 60
Orthogonalité 60
Ostrovski 260

Peano 132
Perpendiculaire 173
Picard 171
Point(s)
fixe 147
ordinaire 262
singulier 262
de Vallron 250
Polynômes
d'Hermite 216
de Jacobi 216
de Laguerre 216
de Legendre 211
de Tchébychev 216
Presque-solution 170

Principe du point fixe 147
Problème
des isopérimètres 195
de Watson 253
Produit
cartésien 16
de convolution 238
d'un opérateur par un nombre 17
Projection d'un vecteur sur un sous-espace 173

Rayon de courbure 271
Rayon vecteur 261
Réseau linéaire 68
Rtesz F 133
Rodrigues 216

Schmidt 133
Schwartz L 77
Série
de Fourier 175
de Fourier-Legendre 214
de vecteurs 50
Serret 287
Sobolev 77
Solution
de l'équation différentielle 134
générale 134
particulière 134
Somme
directe 14
d'opérateurs 17
Sous-algèbre 26
Sous-espace 14
invariant 18
osculateur 265
propre 19
Spectre 27
d'un élément de l'algèbre 116
d'un opérateur linéaire 121
symétrique 29
Sphère osculatrice 287
Spirale de Wiener 285
Stone 132, 133
Suite(s)
convergente 31
en forme de delta 73
Suite d'opérateurs
convergente 102
fortement convergente 103
Système orthonormé 62

Tait 172
Tangente 262

Théorème
d'Arzelà 35
de Banach 99
de Banach-Steinhaus 103
de Broudno 113
de Carleman-Ostrovski 251
de Carleson 207
de Fejér 207
de Gelfand-Mazur 117
de Jackson 82
de Nikolski 209
de Pythagore 80
de Riesz 50
de Robinson 114
de Stone 71
de Toeplitz 111
de Weierstrass 71, 72
Thomson 172
Toeplitz 133
Torsion 273

Transformée
de Fourier 220
de Laplace 242

Unité d'une algèbre 25

Valeur propre 19
généralisée 121
Van der Pol 260
Vecteur courbure 269
Vecteur nul 10
Vecteur propre 19
Vecteurs 10
Volterra 132
Voronoï 133

Weierstrass 132
Wiener 133, 287
Wronski 171

Table des matières

AVANT-PROPOS 5

TROISIÈME PARTIE

CHAPITRES CHOISIS DE L'ANALYSE MODERNE

CHAPITRE 12. STRUCTURES FONDAMENTALES DE L'ANALYSE 9

§ 12.1. Espaces vectoriels 10
§ 12.2. Espaces métriques 31
§ 12.3. Espaces vectoriels normés 41
§ 12.4. Espaces hilbertiens 56
§ 12.5. Approximations dans l'espace des fonctions continues sur un compact 68
§ 12.6. Dérivation et intégration de fonctions à valeurs dans un espace normé 81
§ 12.7. Opérateurs linéaires continus 94
§ 12.8. Algèbres normées 114
§ 12.9. Propriétés spectrales des opérateurs linéaires 121
Exercices 130
Historique 132

CHAPITRE 13. ÉQUATIONS DIFFÉRENTIELLES 134

§ 13.1. Définitions et exemples 134
§ 13.2. Théorème du point fixe 147
§ 13.3. Existence et unicité de la solution d'une équation différentielle dans un espace normé 149
§ 13.4. Système d'équations vectorielles 154
§ 13.5. Equation vectorielle d'ordre supérieur 157
§ 13.6. Equations et systèmes linéaires 159
§ 13.7. Opérateur résolvant d'une équation linéaire homogène 163
§ 13.8. Résolution d'une équation linéaire non homogène . . . 166
Exercices 169
Historique 171

CHAPITRE 14. DÉVELOPPEMENTS ORTHOGONAUX 172

§ 14.1. Développements orthogonaux dans un espace hilbertien 172
§ 14.2. Séries de Fourier classiques 176
§ 14.3. Convergence d'une série de Fourier en un point et sur un ensemble 182
§ 14.4. Autres propriétés des séries de Fourier. Applications . . 191

§ 14.5. Divergence des séries de Fourier et sommation généralisée 205
§ 14.6. Exemples de systèmes orthogonaux 211
Exercices 216
Historique 218

CHAPITRE 15. TRANSFORMATION DE FOURIER 219

§ 15.1. Intégrale de Fourier et son inversion 219
§ 15.2. Autres propriétés de l'intégrale de Fourier 226
§ 15.3. Exemples et applications 238
§ 15.4. Transformation de Laplace 241
§ 15.5*. Classes quasi analytiques de fonctions 249
Exercices 259
Historique 260

CHAPITRE 16. COURBES GAUCHES 261

§ 16.1. Définitions de base 261
§ 16.2. Courbure. Courbures d'ordre supérieur 269
§ 16.3. Dégénérescence de la base naturelle 275
§ 16.4. Equations naturelles 276
§ 16.5. Hélices 282
Exercices 286
Historique 287

INDICATIONS ET RÉPONSES 288
Bibliographie 293
Index 295

A NOS LECTEURS

Les Editions Mir vous seraient très reconnaissantes de bien vouloir leur communiquer votre opinion sur le contenu de ce livre, sa traduction et sa présentation, ainsi que toute autre suggestion.

Notre adresse: Editions Mir, 2, Pervi Rijski péréoulok, Moscou, I-110, GSP, U.R.S.S.

Imprimé en Union Soviétique

A. Ourmaev

LES FONDEMENTS DE LA SIMULATION SUR LES CALCULATEURS ANALOGIQUES

Alexandre Ourmaev, candidat au doctorat ès sciences techniques, examine la programmation des calculateurs analogiques pour la résolution de problèmes techniques et scientifiques. Cet ouvrage expose les principes de fonctionnement des calculateurs analogiques et des principaux blocs opérationnels; on y trouve les méthodes générales de programmation des calculateurs analogiques, ainsi que les méthodes de résolution de classes concrètes de problèmes embrassant plusieurs domaines: mécanique générale, physique, électronique, chimie physique, résistance des matériaux, programmation linéaire, etc. Ce manuel est illustré d'un grand nombre d'exemples ainsi que de problèmes dont la résolution permettra au lecteur d'asseoir ses connaissances théoriques.

Cet ouvrage s'adresse aux étudiants des instituts techniques. Il peut être utilisé avec le plus grand profit par les professeurs dans la préparation de leurs cours sur les applications des calculateurs analogiques. Il rendra également le plus grand service aux ingénieurs et techniciens qui utilisent les calculateurs analogiques.

V. Karmanov

PROGRAMMATION MATHÉMATIQUE

Cet ouvrage a pour but de donner une idée des modèles de gestion, de planification, d'analyse de situations, et des problèmes d'optimisation qui en résultent; il développe également les éléments de la théorie de la programmation mathématique et expose enfin les méthodes fondamentales permettant de résoudre les problèmes de programmation mathématique.

Il est destiné aux élèves des universités et des grandes écoles qui sont déjà initiés aux éléments des mathématiques supérieures. Il intéressera également les personnes désireuses d'approfondir leurs connaissances de la théorie et des méthodes de résolution des problèmes de programmation mathématique.

M. Krasnov, A. Kissélev et G. Makarenko

ÉQUATIONS INTÉGRALES

Le présent Cours est destiné aux étudiants des grandes écoles ayant une formation mathématique poussée et à tous ceux qui s' intéressent aux méthodes de résolution des principaux types d'équations intégrales. On donne au début de chaque chapitre un résumé des principaux résultats, des formules et l'on y étudie en détail des exemples types.